



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 173140

TO: Sheridan Swope
Location: REM-2B71/3C70
Art Unit: 1656
Thursday, December 08, 2005

Case Serial Number: 10/803530

From: Alex Waclawiw
Location: Biotech-Chem Library
Rem 1A71
Phone: 272-2534

Alexandra.waclawiw@uspto.gov

Search Notes

THIS PAGE BLANK (USPTO)

173140

STIC-Biotech/ChemLib

From: Swope, Sheridan
Sent: Friday, December 02, 2005 12:03 AM
To: STIC-Biotech/ChemLib
Subject: 10/803530

For 10/803530, pls search and interference search:

SID 2 against the NT and AA data bases.

Sheridan Swope, Ph.D.
Patent Examiner, AU 1656
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E03C70 Remsen Bld (Mailbox)

Not MacLaurin

Point of Contact:
Alexandra Waclawiw

***** Technical Info. Specialist

***** CM 8A02 Tel 308-448

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: 12-5
Date completed: 12-8
Searcher Prep Time: 8
Online Time: 9

***** Type of Search

NA# _____ AA# 2
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

***** Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: /
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 20:10:51 : Search time 6338 Seconds
(without alignment)
3794.862 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKPLRPK.....VYTKVSALYLNWYVWKAEL 435

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 79147668 seqs, 27645789525 residues

Total number of hits satisfying chosen parameters: 158295336

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame+p2n.model -US10803530/runat 05122005 083217 2181/app query.fasta_1.583
-Q=/cgn2_1/USPTO.spool_p/US10803530 -QFMT=fastap -SUFFIX=rpnm -MINMATCH=0.1 -LOOPCL=0
-DB=Pending Patents NA_Main -MATRIX=blosum62 -TRANS=human40.cdi
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530 @cgn2_1 11077 @runat 05122005 083217 2181 -NCPUS=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA_Main:
1: /cgn2_6/ptodata/1/pna/PCTUSA_COMB.seq:
2: /cgn2_6/ptodata/1/pna/PCTUSB_COMB.seq:
3: /cgn2_6/ptodata/1/pna/PCTUSC_COMB.seq:
4: /cgn2_6/ptodata/1/pna/US06_COMB.seq:
5: /cgn2_6/ptodata/1/pna/US075_COMB.seq:
6: /cgn2_6/ptodata/1/pna/US076_COMB.seq:
7: /cgn2_6/ptodata/1/pna/US077_COMB.seq:
8: /cgn2_6/ptodata/1/pna/US078_COMB.seq:
9: /cgn2_6/ptodata/1/pna/US079_COMB.seq:
10: /cgn2_6/ptodata/1/pna/US080_COMB.seq:
11: /cgn2_6/ptodata/1/pna/US081_COMB.seq:
12: /cgn2_6/ptodata/1/pna/US082_COMB.seq:
13: /cgn2_6/ptodata/1/pna/US083_COMB.seq:
14: /cgn2_6/ptodata/1/pna/US084_COMB.seq:
15: /cgn2_6/ptodata/1/pna/US085_COMB.seq:
16: /cgn2_6/ptodata/1/pna/US086_COMB.seq:
17: /cgn2_6/ptodata/1/pna/US087_COMB.seq:
18: /cgn2_6/ptodata/1/pna/US088_COMB.seq:
19: /cgn2_6/ptodata/1/pna/US089_COMB.seq:
20: /cgn2_6/ptodata/1/pna/US090_COMB.seq:
21: /cgn2_6/ptodata/1/pna/US091_COMB.seq:
22: /cgn2_6/ptodata/1/pna/US092_COMB.seq:
23: /cgn2_6/ptodata/1/pna/US093_COMB.seq:
24: /cgn2_6/ptodata/1/pna/US094_COMB.seq:
25: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:
26: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:
27: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:
28: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:

29: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:
30: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:
33: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:
34: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:
35: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:
36: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:
37: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:
38: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:
39: /cgn2_6/ptodata/1/pna/US099E_COMB.seq:
40: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:
41: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:
42: /cgn2_6/ptodata/1/pna/US101_COMB.seq:
43: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:
44: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:
45: /cgn2_6/ptodata/1/pna/US103A_COMB.seq:
46: /cgn2_6/ptodata/1/pna/US103B_COMB.seq:
47: /cgn2_6/ptodata/1/pna/US103C_COMB.seq:
48: /cgn2_6/ptodata/1/pna/US103D_COMB.seq:
49: /cgn2_6/ptodata/1/pna/US103E_COMB.seq:
50: /cgn2_6/ptodata/1/pna/US103F_COMB.seq:
51: /cgn2_6/ptodata/1/pna/US104_COMB.seq:
52: /cgn2_6/ptodata/1/pna/US105_COMB.seq:
53: /cgn2_6/ptodata/1/pna/US106A_COMB.seq:
54: /cgn2_6/ptodata/1/pna/US106B_COMB.seq:
55: /cgn2_6/ptodata/1/pna/US107A_COMB.seq:
56: /cgn2_6/ptodata/1/pna/US107B_COMB.seq:
57: /cgn2_6/ptodata/1/pna/US107C_COMB.seq:
58: /cgn2_6/ptodata/1/pna/US107D_COMB.seq:
59: /cgn2_6/ptodata/1/pna/US107E_COMB.seq:
60: /cgn2_6/ptodata/1/pna/US107F_COMB.seq:
61: /cgn2_6/ptodata/1/pna/US107G_COMB.seq:
62: /cgn2_6/ptodata/1/pna/US108_COMB.seq:
63: /cgn2_6/ptodata/1/pna/US109A_COMB.seq:
64: /cgn2_6/ptodata/1/pna/US109B_COMB.seq:
65: /cgn2_6/ptodata/1/pna/US109C_COMB.seq:
66: /cgn2_6/ptodata/1/pna/US110A_COMB.seq:
67: /cgn2_6/ptodata/1/pna/US110B_COMB.seq:
68: /cgn2_6/ptodata/1/pna/US110C_COMB.seq:
69: /cgn2_6/ptodata/1/pna/US110D_COMB.seq:
70: /cgn2_6/ptodata/1/pna/US111A_COMB.seq:
71: /cgn2_6/ptodata/1/pna/US111B_COMB.seq:
72: /cgn2_6/ptodata/1/pna/US112_COMB.seq:
73: /cgn2_6/ptodata/1/pna/US117_COMB.seq:
74: /cgn2_6/ptodata/1/pna/US600_COMB.seq:
75: /cgn2_6/ptodata/1/pna/US601_COMB.seq:
76: /cgn2_6/ptodata/1/pna/US602A_COMB.seq:
77: /cgn2_6/ptodata/1/pna/US602B_COMB.seq:
78: /cgn2_6/ptodata/1/pna/US603_COMB.seq:
79: /cgn2_6/ptodata/1/pna/US604A_COMB.seq:
80: /cgn2_6/ptodata/1/pna/US604B_COMB.seq:
81: /cgn2_6/ptodata/1/pna/US605_COMB.seq:
82: /cgn2_6/ptodata/1/pna/US606_COMB.seq:
83: /cgn2_6/ptodata/1/pna/US607_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	1305	40	US-10-030-688-1
2	2342	100.0	2154	28	US-09-606-680-3113
3	2338	99.8	2038	29	US-09-659-151-18
4	2338	99.8	2038	42	US-10-180-719-18
5	2338	99.8	2038	66	US-11-045-577-18
6	2338	99.8	2038	71	US-11-183-914-18
7	2337	99.8	1314	41	PCT-US02-19257-88
8	2337	99.8	1314	42	US-10-126-052A-448

Sequence 1, Appli
Sequence 3113, Ap
Sequence 18, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 88, Appl
Sequence 448, Appl

```
9 2337 99.8 1314 42 US-10-173-999-88 Sequence 88, Appl
10 2337 99.8 1314 44 US-10-295-027-133 Sequence 133, App
11 2337 99.8 1314 44 US-10-295-027-778 Sequence 778, App
12 2337 99.8 1314 44 US-10-295-027-790 Sequence 790, App
13 2337 99.8 1314 44 US-10-295-027-830 Sequence 830, App
14 2337 99.8 1314 44 US-10-295-027-979 Sequence 979, App
15 2337 99.8 2079 82 US-60-625-561-448 Sequence 448, App
16 2337 99.8 2104 3 PCT-US04-21227-1 Sequence 1, Appli
17 2337 99.8 2104 64 US-10-956-157-2292 Sequence 2, Appli
18 2337 99.8 2104 65 US-10-991-287-2 Sequence 2, Appli
19 2337 99.8 2104 81 US-60-507-511-2292 Sequence 2, Appli
20 2337 99.8 2307 1 PCT-US02-07826-317 Sequence 2292, Ap
21 2337 99.8 2307 1 PCT-US02-18638A-217 Sequence 317, App
22 2337 99.8 2307 40 US-10-097-340-317 Sequence 317, App
23 2337 99.8 2307 42 US-10-171-311-217 Sequence 217, App
24 2337 99.8 2307 66 US-10-050-926-317 Sequence 317, App
25 2337 99.8 2121 1 PCT-US01-18568-1 Sequence 1, Appli
26 2337 99.7 2121 1 PCT-US01-18568A-1 Sequence 1, Appli
27 2335 99.7 2121 62 US-10-803-530-1 Sequence 1, Appli
28 2335 99.7 2121 51 US-10-417-375-139 Sequence 139, App
29 2335 99.6 2590 51 US-10-417-375A-139 Sequence 139, App
30 2333 99.6 2590 51 US-10-417-375B-139 Sequence 139, App
31 2333 99.6 2590 82 US-60-625-561-447 Sequence 447, App
32 2333 99.6 2590 82 US-10-030-688-3 Sequence 3, Appli
33 2333 99.4 2165 1 PCT-US02-08456-634 Sequence 634, App
34 2329 99.4 2165 42 US-10-101-510-634 Sequence 634, App
35 2329 99.4 2165 71 US-11-146-198-634 Sequence 634, App
36 2329 99.4 2108 3 PCT-US05-22501-4335 Sequence 4335, Ap
37 2328 99.4 2112 42 US-10-170-235-14349 Sequence 14349, A
38 2328 99.4 2627 51 US-10-417-375-141 Sequence 141, App
39 2328 99.4 2627 51 US-10-417-375A-141 Sequence 141, App
40 2328 99.4 2627 51 US-10-417-375B-141 Sequence 141, App
41 2328 99.4 2627 82 US-60-625-561-446 Sequence 446, App
42 2324 99.2 2070 3 PCT-US99-19655-2 Sequence 2, Appli
43 2324 99.2 2070 31 US-09-763-978A-2 Sequence 2, Appli
44
45

ALIGNMENTS

RESULT 1
US-10-030-688-1
; Sequence 1, Application US/10030688
; GENERAL INFORMATION:
; APPLICANT: Merck Patent GmbH
; TITLE OF INVENTION: Seripancrin
; FILE REFERENCE: SeripancrinUHS
; CURRENT APPLICATION NUMBER: US/10/030,688
; CURRENT FILING DATE: 2002-01-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1305)
US-10-030-688-1
Alignment Scores:
Pred. No.: 0
Score: 2342.00 Length: 1305
Percent Similarity: 100.00% Matches: 435
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Indels: 0
DB: 40 Gaps: 0

US-10-803-530-2 (1-435) x US-10-030-688-1 (1-1305)

Qy 1 MetAspProAspSerGlnProLeuAsnSerLeuAspValLeuProLeuArgLysPro 20
|||||
```


; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ;
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/659,151
 ; FILING DATE: 11-Sep-2000
 ;
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/008,271
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ;
 ; INFORMATION FOR SEQ ID NO: 18:
 ;
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2038 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNOT13
 ; CLONE: 1337018
 ;
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 ; US-09-659-151-18

Alignment Scores:
 Pred. No.: 0 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 29 Gaps: 0

US-10-803-530-2 (1-435) x US-09-659-151-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspVallyProLeuArgLysPro 20
 Db 200 ATGGATCTGACAGTGATCACTCTGACAGCTCTGATGTCACAAACCCCTGGCGAAACC 259
 Qy 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSer 40
 Db 260 CGTATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGC 319
 Qy 41 LeuAlaSerIleIleValValValValIleLysValIleLeuAspLysTyrPhe 60
 Db 320 CTGGCGAGTATCATCTGTTGTTCTCATCAAGGTGATCTGGATAAATACTATTCT 379
 Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
 Db 380 CTCTGGCGGAGCGCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 439
 Qy 81 CysProLeuGlyGluAspGluGluHisCysVallySerPheProGluGlyProAlaVal 100
 Db 440 TGTCTCTGGGGAGGAGGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499
 Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
 Db 500 GCAGTCGGCTCTCCAGGAGCCGATCCACACTGAGGTGCTGAGCTCGGCCACAGGGAAC 559

Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 Db 560 TGGTTCCTCTGCTGTTTGCACAACCTTCACAGAAGCTCTCGTAGACAGCCTGTAGGCAG 619
 Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
 Db 620 ATGGGCTACAGCAGCAACCCACTTTTCTAGAGCTGTGGAGATTGGCCAGACCAAGATCTG 679
 Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 Db 680 GATGTTGTTGAATTCACAGAAACACAGCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCC 739
 Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 Db 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
 Qy 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 Db 800 CCCGCTGTGGTGGTGGGAGGAGGCTCTGTGATTTCTTGGCTTGGCAGGTGAGCATC 859
 Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
 Db 860 CAGTACGACAAACAGCAGCGCTCTGTGGAGGAGCATCTGGACCCCACTGGGTCTCTCAGC 919
 Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
 Db 920 GCAGCCCACTGCTTCAGGAACATACCGATGTGTTCACTGGAAAGTGGGGCAGGCTCA 979
 Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 Db 980 GACAAACTGGGCAGCTTCCATCCCTGGCTGTGGCCAAAGATCATCATCATTTGATTTCAAC 1039
 Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 Db 1040 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTC 1099
 Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
 Db 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
 Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
 Db 1160 CCACCTCTGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCACATA 1219
 Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
 Db 1220 CTGCTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGCTGCAATGCAGACGATCGTAC 1279
 Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 Db 1280 CAGGGGGAAGTCACCGAGAAGATGATGTGTGCGAGCATCCCGGAAGGGGGTGTGACACC 1339
 Qy 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 Db 1340 TGCCAGGTTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCACTGGCATGTGGTGGGC 1399
 Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
 Db 1400 ATCGTTAGCTGGGCTATGGTCTGGGGGCCCGGAGACCCCGAGGAGTATACCAAGGTC 1459
 Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1460 TCAGGCTATCTCAACTGGATCTCAATGTCTTGGAAAGGCTGAGCTG 1504

RESULT 4

; US-10-180-719-18
 ; Sequence 18, Application US/10180719
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry
 ; Guegler, Karl J.
 ; Corley, Neil C.

; Tang, Tom Y.
 ; Shah, Purvi
 ; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Dr.
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/180,719
 ; FILING DATE: 25-Jun-2002
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/008,271
 ; FILING DATE: 16-Jan-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mohan-Peterson, Sheela
 ; REGISTRATION NUMBER: 41,201
 ; REFERENCE/DOCKET NUMBER: PF-0458 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2038 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: COLNNOT13
 ; CLONE: 1337018
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
 ; US-10-180-719-18

Alignment Scores:

Pred. No.: 0 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 42 Gaps: 0

US-10-803-530-2 (1-435) x US-10-180-719-18 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
 DB 200 ATGGATCCTGACAGTATCAACCTCTGAACAGCTCGATGTCAAAACCCCTGGCAACCC 259
 QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
 DB 260 CGTATCCCCATGAGACCTTCAGAAAGTGGGATCCCCATCATCATAGCACTACTGAGC 319
 QY 41 LeuAlaSerIleIleValValValValLeuIleLysValIleLeuAspLysTyrPhe 60
 DB 320 CTGGCAGTATCATATTGGTGTCTCATCAAGTGATTCGGATAAATACTACTTTC 379
 QY 61 LeuCyGlyGlnProLeuHisPheIleProArgLysGlnLeuCyAspGlyGluLeuAsp 80
 DB 380 CTCGGCGGCGAGCTCTCCACTTCATCCCGAGAGCAGCTGTGTACGAGAGCTGGAC 439
 QY 81 CyAspProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
 DB 440 TGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAGGGCTTGAGTG 499
 QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120

DB 500 GCAGTCCGCTCTCCNAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGAAC 559
 QY 121 TrpPheSerAlaCyAspPheAspPheThrGluAlaLeuAlaGluThrAlaCyAspArgGln 140
 DB 560 TGGTCTCTGCTGCTTTGCAAACTTCACAGAAGCTCTCGCTGACACAGCCCTGTAGGCAG 619
 QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
 DB 620 ATGGGCTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCCAAGACAGGATCTG 679
 QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 DB 680 GATGTTGTTGAATTCACAGAAAACGCGCAGAGCTTCGCATGCGGAATCTCAAGTGGGCC 739
 QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 DB 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGTGCTGTGGGAGAGCCGTGAAGAC 799
 QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 DB 800 CCCCCTGTGGTGGGTGGGAGAGGCCCTCTGTGGATTCTTGGCCCTTGGCAGGTTCAGCATC 859
 QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
 DB 860 CAGTACGACAAACAGCAGCAGTCTGTGGAGGAGCATCTTGGACCCCTCCACTGGGTCTCAG 919
 QY 241 AlaAlaHisCyAspPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
 DB 920 GCAGCCCCACTGCTTCAGGAACATACCGATGTGTTCACCTGGAAGGTGCGGCGAGCTCA 979
 QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 DB 980 GACAAACTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTAATTCAC 1039
 QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 DB 1040 CCCATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCAGTTCACACTCACTTTC 1099
 QY 301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
 DB 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1159
 QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
 DB 1160 CCACCTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATA 1219
 QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
 DB 1220 CTGCTGCAAGCGCTCAGTCCAGGTCAATTGACACACACCGGTGCATGACAGCAGTGGCTAC 1279
 QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 DB 1280 CAGGGGGAAGTCAACGAGAAGATGATGTGTGAGGATCCCGAAGGGGGTGTGGACACC 1339
 QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValGly 400
 DB 1340 TGCCAGGGTGACAGTGGTGGGCCCTGTGTACCAATCTGACCACTGGCAGTGGCTGTGGGG 1399
 QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
 DB 1400 ATCGTGTAGCTGGGGCTATGGCTGCGGGGGGCCGAGCAGCCCCAGGAGTATACACCAAGGC 1459
 QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1460 TCAGCTATCTCACTGGATCTACAAATGTCTGGNAGGCTGAGCTG 1504

RESULT 5

US-11-045-577-18
 ; Sequence 18, Application US/11045577
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; Hillman, Jennifer L.
 ; Yue, Henry

;; Guegler, Karl J.
;; Corley, Neil C.
;; Tang, Tom Y.
;; Shah, Purvi
;; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
;; NUMBER OF SEQUENCES: 24
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Dr.
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/11/045,577
;; FILING DATE: 27-Jan-2005
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US/09/659,151
;; FILING DATE: 11-Sep-2000
;; APPLICATION NUMBER: 09/008,271
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mohan-Peterson, Sheela
;; REGISTRATION NUMBER: 41,201
;; REFERENCE/DOCKET NUMBER: PF-0458 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2038 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: COLNNOT13
;; CLONE: 1337018
;; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-11-045-577-18

Alignment Scores:
Pred. No.: 0 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 66 Gaps: 0

US-10-803-530-2 (1-435) x US-11-045-577-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db ATGGATCTGACATGATCACTCTGACAGCCCTCGATGCAACCCCTGGCGAAACCC 259
Qy 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleLeuLeuLeuSer 40
Db CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGC 319
Qy 41 LeuAlaSerIleIleValValValValValValValValValValValValValVal 60
Db CTGGCGAGTATCATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db CTCTCGGGCAGCCTCTCCACTTCACTCCAGGAGCAGCTGTGACGGAGAGCTGGAC 439
Qy 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100

Db TGTCCCTTGGGGGAGGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499
Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db GCAGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGNAAC 559
Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db TGGTCTCTGCTGCTGTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCCTGTAGGCAG 619
Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db ATGGGCTACAGCAGCAAAACCACTTTCAGAGCTGTGGAGATTGGCCAGACCAAGATCTG 679
Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db GATGTTGTTGAAATTCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCC 739
Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGGCTTGGCGAGAGCCTGAAGACC 799
Qy 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db CCCCGTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGCATC 859
Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db CAGTAGCAGCAAAACAGCAGCTCTGTGGAGGAGCATCTGGACCCCACTGGGCTCTCACG 919
Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db GACAACTGGGCGAGCTTCCCATCCCTGGTGTGGCCCAAGATCATCATATTGAATTCAC 979
Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db GACAACTGGGCGAGCTTCCCATCCCTGGTGTGGCCCAAGATCATCATATTGAATTCAC 1039
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db CCATATGCCCAAGACAAATGATCGCCCTCATGAAGCTGAGTCCAGTCCACTCTCTTC 1099
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTAGAGAGCTCACTCCAGCCACC 1159
Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db CCACCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGAAGATGTCTGACATA 1219
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db CTGCTGAGGCGTCAGTCCAGGTCAATTGACACACACCGTGCATGCAATGACAGCATGCTAC 1279
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db CAGGGGGAAGTCACCGAGAAGATGATGTGTGAGGCATCCCGGAAGGGGTGTGGACACC 1339
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db TGCCAGGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCTGAGTGTGGTGGGC 1399
Qy 401 IleValSerTrpGlyTyrCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db ATCGTTAGCTGGGCTATGGCTGGGGGCCCCCTGATGTACCAATCTGACCTGAGTGTGGTGG 1459
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db TCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 6

US-11-183-914-18

; Sequence 18, Application US/11183914

GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Hillman, Jennifer L.
 APPLICANT: Yue, Henry
 APPLICANT: Guegler, Karl J.
 APPLICANT: Corley, Neil C.
 APPLICANT: Tang, Tom Y.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Dr.
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/11/183,914
 FILING DATE: 19-JULY-2005
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/09/008,271
 FILING DATE: 16-Jan-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: <Unknown>
 FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
 NAME: Mohan-Peterson, Sheela
 REGISTRATION NUMBER: 41,201
 REFERENCE/DOCKET NUMBER: PF-0458 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-855-0555
 TELEFAX: 650-845-4166
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:

LENGTH: 2038 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: COLNOT13
 CLONE: 1337018
 US-11-183-914-18

Alignment Scores:
 Pred. No.: 0 Length: 2038
 Score: 2338.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.77% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 71 Gaps: 0

US-10-803-530-2 (1-435) x US-11-183-914-18 (1-2038)

QY 1 MetAppProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysPro 20
 DB 200 ATGGATCCTGACAGTGCATCAACCTCTGAAACAGCCTCGATGTCAACCCCTGGCAACCC 259
 QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleIleAlaLeuSer 40
 DB 260 CGTATCCCCATGAGACCTTCAGAAAGGTGGGATCCCCATCATCATGACTACTGAGC 319
 QY 41 LeuAlaSerIleIleIleValValValLeuIleValLeuValIleLeuAspLysTyrPhe 60
 DB 320 CTGGCGAGTATCATCATTTGGTGTTCCTCATCAAGTGATTTCTGGATAAATACTACTTC 379
 QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80

DB 380 CTCGGCGGACGCTCTCCACTTCCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 439
 QY 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
 DB 440 TGTCCCTTGGGGAGGACGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTG 499
 QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAen 120
 DB 500 GCAGTCCCGCTCTCCAGGACCGATCCACACTGCGAGGTGCTGGACTCGGCGACAGCCGAGAAC 559
 QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 DB 560 TGGTCTCTGCTGTTTCGACACACTTCACAGAAAGCTCTCGCTGACAGAGCTGTAGGCAG 619
 QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
 DB 620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGGCCACAGACGAGTCTG 679
 QY 161 AspValValGluIleThrGluAenSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
 DB 680 GATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCC 739
 QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
 DB 740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGGCTGTGGGAGAGCCCTGAAGACC 799
 QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
 DB 800 CCCCCTGTGGTGGGAGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGCATC 859
 QY 221 GlnTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThr 240
 DB 860 CAGTACGACAAACAGCACGCTCTGTGGAGGAGCATCTTGGACCCCATCTGGTCTCTCAG 919
 QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
 DB 920 GCAGCCCACTGCTTCAGAAACATACCATGTTCACTGGAAGGTGGCGGACGCTCA 979
 QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
 DB 980 GACAAACTGGGCAGCTTCCCATCTGCTGTGGCCAGATCATCATCATTAATTCAC 1039
 QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
 DB 1040 CCCATGTATCCCAACAGCAATGACATCGCCCTCATGAAAGTGCAGTCTCCACTCATCTTC 1099
 QY 301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
 DB 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGCTCATCTCAGCCACC 1159
 QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
 DB 1160 CCACCTCTGATCATTTGGATGGGCTTTACGAGAGCAAGATGGAGGAGAGATGTCCTGACATA 1219
 QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
 DB 1220 CTGCTGCAGGCGTCAGTCCAGGTCAATGACAGCACACGTCGAATGACAGCATGGGTAC 1279
 QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValLeuAspThr 380
 DB 1280 CAGGGGGAAGTCCACGAGAAGATGATGTGTGAGGATCTCCCGAAGGGGTGTGGACACC 1339
 QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 DB 1340 TGCCAGGTGACAGTGTGGGCCCTGATGATACCAATCTGACAGTGGCATGTGGTGGGC 1399
 QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
 DB 1400 ATCGTGTAGTGGGCTATGGCTGCGGGGCCCGAGCACCCAGGAGTATACACCAAGGTC 1459
 QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1460 TCAGGCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 7

PCT-US02-19297-88
 ; Sequence 88, Application PC/TUS0219297
 ; GENERAL INFORMATION:
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Eos Biotechnology Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
 ; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
 ; FILE OF INVENTION: Cancer
 ; FILE REFERENCE: 018501-002420PC
 ; CURRENT APPLICATION NUMBER: PCT/US02/19297
 ; CURRENT FILING DATE: 2002-06-18
 ; PRIOR APPLICATION NUMBER: US 60/299,234
 ; PRIOR FILING DATE: 2001-06-18
 ; PRIOR APPLICATION NUMBER: US 60/315,287
 ; PRIOR FILING DATE: 2001-08-27
 ; PRIOR APPLICATION NUMBER: US 60/317,544
 ; PRIOR FILING DATE: 2001-09-05
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/372,246
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 164
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 88
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 PCT-US02-19297-88

Alignment Scores:

Pred. No.: 0 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 1 Gaps: 0

US-10-803-530-2 (1-435) x PCT-US02-19297-88 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
 Db 10 GATCTGACGTGATCAACCTCTGACAGCCTCGATGTCAAAACCCCTGCGCAACCCCGT 69
 QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleIleAlaLeuSerLeu 41
 Db 70 ATCCCCATGGAGACCTTCAGAAAGTGGGATCCCCATCATCATAGCACTACTGAGCCTG 129
 QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
 Db 130 CGAGATATCATGTGTGTGTCTCATCAAGGTGATCTCGGATAAATACTACTTCTCTC 189
 QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
 Db 190 TCGGGCAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT 249
 QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 Db 250 CCCTTGGGGAGGAGCAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCTCGCAGTGGCA 309
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 Db 310 GTCCGCTCTCCAGGAGCCAGTCCACACTCAGGTGCTGAGCTCGGCCACAGGGAACCTGG 369
 QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTCTGCTGTTTCGACAACCTTCACAGAGCTCTCGCTGAGACAGCCTGTAGCGAGATG 429
 QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACAGCAGCAACCCACTTTTCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGAT 489

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGTAATCAAGAAACAGCCAGAGCTTCGCATGCGGAACTCAAGTGGGCGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysAlaLeuAlaCysGlyLysSerLeuLysThrPro 201
 Db 550 CTCTCAGGCTCCCTGGTCTCCCTGCATGTCTGCCCTGTGGGAAGAGCCTGAACACCCCC 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGGAGGCTCTGTGGATCTTCTTGGCTTGGCAGGTCAGATCCAG 669
 QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
 Db 670 TACGACAAACAGCACGCTCTGTGGAGGAGCATCTCGACCCCACTGGTCTCTCACGGCA 729
 QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 Db 730 GCCCACTGCTTCAGGAACATACCGATGTGTCACTGGAGAGGTGCGGCGAGGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
 Db 790 AAATCGGCGAGCTTCCCATCCCTGGCTGTGCCAAGATCATCATTTGAATTCAACCCC 849
 QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCCAACAGCAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 Db 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGGAGGCTCACTCCAGGCCACCCCA 969
 QY 322 LeuTrpIleIleIleIleTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 Db 970 CTCTGGATCATTTGGATGGGCTTTAGCAAGCAGAAATGGAGGAAGATGTCTGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrAtqCysAsnAlaAspAlaTyrGln 361
 Db 1030 CTGCAGGCTCAGTCCAGGTTCATTGACAGCACAGGTGCAATGCAGACCATGCGTACCAG 1089
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 Db 1090 GGGGAAGTCACGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1150 CAGGGTGACAGTGTGGGCCCCCTGTATGTACCAATCTGACCACTGTGGTCATGTGGTGGCATC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 Db 1210 GTTAGCTGGGGCTATGGCTCGGGGGCCCGACACCCCAAGAGTATACCAAGGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 Db 1270 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311

RESULT 8

US-10-126-052A-448
 ; Sequence 448, Application US/10126052A
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Lung Cancer, Compositions and
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
 ; FILE OF INVENTION: Methods of Screening for Modulators of Lung Cancer
 ; FILE REFERENCE: 018501-001530US
 ; CURRENT APPLICATION NUMBER: US/10/126,052A
 ; CURRENT FILING DATE: 2002-04-18
 ; PRIOR APPLICATION NUMBER: US 60/284,770
 ; PRIOR FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: US 60/290,492
 ; PRIOR FILING DATE: 2001-05-10

; PRIOR APPLICATION NUMBER: US 60/339,245
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/334,370
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 691
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 448
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-126-052A-448

Alignment Scores:
Pred. No.: 0 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 42 Gaps: 0

US-10-803-530-2 (1-435) x US-10-126-052A-448 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
DB 10 GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCGAACCCTCGT 69
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 70 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATGACACTACTGAGCCTG 129
QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
DB 130 GCGAGTATCATCATTTGTGTCTCATCAAGGTGATTTCTGGATAAATACTACTCTTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyLeuLeuAspCys 81
DB 190 TCGGGGACGCTCTCCACTTCATCCGAGAGAGCAGCTGTGTGACGAGAGCTGGACTGT 249
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
DB 250 CCCTTTGGGGAGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGGCA 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAenTrp 121
DB 310 GTCCGCTCTCAAGAGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGAACTGG 369
QY 122 PheSerAlaCysPheAspAenPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 370 TTCTCTGCTGTTCGACAACTTCACAGAGCTCTCGCTGACAGCCTGTAGGCAGATG 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 430 GGCTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCACAGCAGGATCTGGAT 489
QY 162 ValValGluIleThrGluAenSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 490 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACCTCAAGTGGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 550 CTCTCAGGCTCCCTGTGTCTCCCTGCACTGTCTTGTGCTGTGGGAAGAGCCTGAGACCC 609
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 610 CGTGTGGTGGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGATCCAG 669
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 670 TACGACAAACAGCACGCTCTGTGGAGGAGCATCTCTGAGCCCTCCACTCTGGTCTCTCACGGCA 729

QY 242 AlaHisCysPheArgLysHisThrAspValPheAenTrpLysValAlaCysArgSerAsp 261
DB 730 GCCCACTGCTTCAGGAAACATACCGATGTGTTCACCTGGAAGGTGCGGCGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
DB 790 AACTGGGCGCTTCCCATCCCTGGCTGTGGCCAGATCATCATCATTAATTCAACCCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAGACAATGACATGCCCTCATGAAGCTGCAGTTCCTCCACTCCTTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
DB 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 969
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAenGlyGlyLysMetSerAspIleLeu 341
DB 970 CTCTGGATCATGGATGGGCTTTACGAAGCAGAAATGGAGGAGATGTCTGACATCTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
DB 1030 CTGCAGGCGTCAGTCCAGTCAATTGACAGCACACCGTGCATTCACAGCATCGTACCCAG 1089
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1090 GGGGAAGTCACCGAGAAGATGTGTGACGACATCCCGAAGGGGTGTGGACACCTGC 1149
QY 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
DB 1150 CAGGTGACAGTGTGGGCCCCCTGATGTACCAATCTGACCATGTGTCATGTGGGCGATC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
DB 1210 GTTACCTGGGCTATGGCTGCGGGGCGCCGAGCACCCAGGAGTATACACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAenTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1270 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311

RESULT 9

US-10-173-999-88
; Sequence 88, Application US/10173999
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 88
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-999-88

Alignment Scores: 0 Length: 1314
Pred. No.: 2337.00 Matches: 434
Score:

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	42	Gaps:	0
US-10-803-530-2 (1-435) x US-10-173-999-88 (1-1314)			
QY	2	AspProhSerSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg	21
Db	10	GATCCTTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGGCGAAACCCCGT	69
QY	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu	41
Db	70	ATCCCATGGAGACCTTCAGAAGGTGGGATCCCATCATCATGAGCACTACTGAGCGCTG	129
QY	42	AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu	61
Db	130	GCAGATATCATATTGTGGTTGTCTCATCAAGGTGATCTTGGATAATACTACTTCCTC	189
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	190	TGGCGCAGCCTTCACATTCCCGAGGAAGCAGCTGTGTGACGAGAGCTGGACTGT	249
QY	82	ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	250	CCCTTGGGGAGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTGCGCA	309
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	310	GTCCGCGCTTCCAAGGACCGATCCACATGCGAGGTGTGGACTCGGCCACAGGAACTGG	369
QY	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	370	TTCTCTGCTGTTCGACAACTTCACAGAGCTCTCGCTGAGACAGCGCTGTGAGCGATG	429
QY	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	430	GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTGGCCACAGCAGGATCTGGAT	489
QY	162	ValValGluIleThrGluAsnSerGlyGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	490	GTGTGTGAATCACAGAAACAGCAGAGCTTCGCATGCGAACTCAAGTGGCGCCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	550	CTCTCAGGCTCCCTGGTCTCCCTGCTGCTTGGCTGTGGGAAGAGCTTGAGACCCCC	609
QY	202	ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	610	CGTGTGGTGGGTGGGAGGCGCTCTGTGGATTCTTGGCTTTGGCAGGTGCAGCATCCAG	669
QY	222	TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	670	TACGACAAACAGCAGCTCTGTGGAGGAGGATCTCTGACCCCCCATCTGGGTCTCTACGGCA	729
QY	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	730	GCCCACTGCTTCAGGAACAATACCGATGTGTTCAACTGGAAGGTGGCGGCGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro	281
Db	790	AAACTGGGCAGCTTCCCATCCCTGGCTGTGGCCCAAGATCATCATATTGAATTCACCCCC	849
QY	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTACCCCAAGACAATGATCATGCGCCTCATGAAGCTGCAGTTCCTCATCTCTTCTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGluLeuThrProAlaThrPro	321
Db	910	GGCACAGTCAGGCCCATCTGTCTGCCCTCTTTGTATGAGAGCTCACTCCAGCACCCCA	969
QY	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.73% Indels: 0
DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-133 (1-1314)

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 10 GATCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCTCTGGCAAAACCCG 69
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTAGACCTG 129
Qy 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrrPheLeu 61
Db 130 GCGAGTATCATCATTTGTGTCTCTCATCAAGGTGATTCGTGATAAATACTACTTCCTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TCGGGGAGCCCTCTCATCTTCACTCCGAGAGCAGCTGTGTGACGGAGAGCTGGACTGT 249
Qy 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 250 CCCTTGGGGAGGAGCGAGGACACTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGCA 309
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCCTCTCCAAAGCGATCCACACTGACAGTGTGGACTCGGACACAGGAACTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTGTTCACAACTTCACAGAACTCTCGCTGACAGACCCCTGTAGGAGATG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATGGCCAGACCAAGATCTGGAT 489
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCCCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCTCAGGCTCCCTGCTCTCCCTGCTGCTGTTCCTGTGGAAAGAGCCTGAAGACCC 609
Qy 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGTGGGAGAGGCTCTGTGGATTCCTGGCCCTGGCAGGTCAGCATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TAGCAAAACAGCACGCTGTGGAGGAGGATCTCGACCCCTCCACTGGGTCTCTCAGGCA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAGGAAACATACCGATGTGTTCACTGGAAAGTGGGGCAGGCTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleLeuPheAsnPro 281
Db 790 AAATCGGCAGCTTCCATCCCTCGCTGTGGCCAGATCATCATCATTAATTCACACCC 849
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTTCCTCACTTCTCA 909
Qy 302 GlyThrValArgProIleCysValLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATCTGTCTGCCCCCTTTTGTGAGGAGGCTCACTCCAGCACCCCA 969
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
```

```
Db 970 CTCTGGATCATTTGGATGGGCTTTTACGAGCAGAAATGGAGGAAAGATCTCTGACATCTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1030 CTGAGGCGTCAGTCCAGGTCTATTGACAGCACACGGTGTCAATGCAGACGATGCGTACCAG 1089
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1090 GGGGAAGTCACCGAGAGATGATGTGTGAGGCATCCCGGAAGGGGGTGTGGACACCTGC 1149
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACAGTGCATGTGGTGGGCATC 1209
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSer 421
Db 1210 GTTAGCTGGGCTATGGCTGGGGGGCCCGAGCACCCCGAGGATATACCAAGGTCTCA 1269
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311
```

RESULT 11

US-10-295-027-778

; Sequence 778, Application US/10295027

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

; FILE REFERENCE: 018501-012500US

; CURRENT APPLICATION NUMBER: US/10/295,027

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: US 09/663,733

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13

; PRIOR APPLICATION NUMBER: US 60/335,394

; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464

; PRIOR FILING DATE: 2001-11-21

; PRIOR APPLICATION NUMBER: US 60/334,393

; PRIOR FILING DATE: 2001-11-29

; PRIOR APPLICATION NUMBER: US 60/340,376

; PRIOR FILING DATE: 2001-12-14

; PRIOR APPLICATION NUMBER: US 60/347,211

; PRIOR FILING DATE: 2002-01-08

; PRIOR APPLICATION NUMBER: US 60/347,349

; PRIOR FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/355,250

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: US 60/356,714

; PRIOR FILING DATE: 2002-02-13

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 1386

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 778

; LENGTH: 1314

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-295-027-778

Alignment Scores:

Pred. No.: 0

Score: 2337.00

Length: 1314

Matches: 434

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	44	Gaps:	0
US-10-803-530-2 (1-435) x US-10-295-027-778 (1-1314)			
QY	2	AspProAspSerAspGlnProLeuAenSerLeuAaspVallysProLeuArlgVysProArg	21
DB	10	GATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAACCCCTCGGCAACCCCGT	69
QY	22	IleProMetGluThrPheArgLysValGlyLeProIlellelleAlaLeuLeuSerIleu	41
DB	70	ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATAGCACTACTGAGCCTG	129
QY	42	AlaSerIlellelleValValValLeuIleLysValIleLeuApsLysTyTyPheIleu	61
DB	130	CGGAGTATCATATTGTGGTTGTCTTCATCAAGGTGATTCGGATAAATACTACTCTCTC	189
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuApsCys	81
DB	190	TGCGGGCAGCCTCTCCACTTCATCCGAGAGCAAGCTGTGTGACGAGAGCTGGACTGT	249
QY	82	ProLeuGlyGlnAaspGluLuhisCysVallysSerPheProGluGlyProAlaValAla	101
DB	250	CCCTTGGGGGAGACAGAGGAGCACTGTGTCAAGAGCTTCCCAGAAAGGCCCTGCAGTGGCA	309
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAaspSerAlaThrGlyAenTrp	121
DB	310	GTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGGAACCTGG	369
QY	122	PheSerAlaCysPheAspAenPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
DB	370	TTCTCTGCCTGTTTCGACAACTTCACAGAGCTCTCGCTGAGACAGCCTGTAGCGAGATG	429
QY	142	GlyTySerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAaspLeuAasp	161
DB	430	GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGGACTCTGGAT	489
QY	162	ValValGluIleThrGluAenSerGlnGluLeuArgMetArgAenSerSerGlyProCys	181
DB	490	GTTTGTGAAATCACAAACAGCCAGAGAGCTTCGCATGCGGAACTCAAGTGGGGCCCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
DB	550	CTCTCAGGCTCCCTGCTCTCCCTGCACCTCTTGCTGTGGGAAGAGCCTGAAGACCCCC	609
QY	202	ArgValValGlyGlyGlnAlaSerValAaspSerTrpProTrpGlnValSerIleGln	221
DB	610	CGTGTGGTGGGTGGGGAGAGCCCTCTGTGGATTCTTGCCCTTGGCAGGTTCAGCATCCAG	669
QY	222	TyrApsLysGlnHisValCysGlyClySerIleLeuAaspProHisTrpValLeuThrAla	241
DB	670	TACGACAAACAGCAGCTCTGTGGAGGAGCATCTGGACCCCCACCTGGGTCTTCAGCGCA	729
QY	242	AlaHisCysPheArgLysHisThrApsValPheAenTrpLysValArgAlaGlySerAasp	261
DB	730	GCCCACTGCTTCAGGAACATACCGATGTTCAACTGGGAAGGTGGGGAGGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIlellellelleGluPheAenPro	281
DB	790	AAACTGGGCAGCTTCCCATCCCTGTGGTGGCCAAAGATCATCATTTGAATTCAACCCC	849
QY	282	MetTyProLysAspAasnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
DB	850	ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCACACTCACTTCTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAaspGluLeuThrProAlaThrPro	321
DB	910	GGCAGAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGGTCACTCAAGCCACCCA	969
QY	322	LeuTrIlellelleGlyTrpGlyPheThrLysGlnAenGlyLysMetSerAaspIleLeu	341

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	99.79%	Indels:	0
DB:	44	Gaps:	0

US-10-803-530-2 (1-435) x US-10-295-027-790 (1-1314)

Qy	2	AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg	21
Db	10	GATCCTGCAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAAAACCCCGT	69
Qy	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu	41
Db	70	ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTAGACCTG	129
Qy	42	AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu	61
Db	130	CGGAGTATCATATTGTGTGCTCATCAAGGTGATTTCTGGATAAATACTACTTCCCTC	189
Qy	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	190	TGCGGGCAGCCTCTCCACTTCATCCCGAGGAACAGCTGTGTGACGGAGAGCTGACTGT	249
Qy	82	ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	250	CCCTTGGGGAGGACGAGAGACACTGTGTCAAGAGCTTCCCCGAAAGGCGCTCGAGTGGCA	309
Qy	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	310	GTCCGCTCTCCAAGACCGATCCACACTGCAGGTGTGGAATCGGCCACAGGAACCTGG	369
Qy	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	370	TTCTCTGCTGTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATG	429
Qy	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	430	GGCTACAGCAGCAAAACCACTTTCAGAGCTGTGGAGATTGGGCCACAGACATCTGGAT	489
Qy	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	490	GTTGTTGAATACACAGAAAACAGCAGGAGCTTCGATCGGAATCAAGTGGGCCCTGT	549
Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	550	CTCTCAGGCTCCCTGCTCTCCCTGCACCTGTCTTGCTGTGGGAGAGCCTGAAGACCCCC	609
Qy	202	ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	610	CCTGTGGTGGGTGGGGAGAGGAGCCCTCTGTGGGATTTCTTGGCCCTTGGCAGGTACGAC	669
Qy	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla	241
Db	670	TACGCAAAACACGACGTCTGTGGAGGAGCATCCTGGACCCCCACCTGGGTCTCACGGCA	729
Qy	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	730	GCCCACTGCTTCAGGNAACATACCGATGTGTTCACTGGAAAGGTGGCGCAGGCTCAGAC	789
Qy	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro	281
Db	790	AAACTGGGCAGCTTCCATCCCTGGCTGTGGCGGACATCATCATCATTTGAATTCACACCC	849
Qy	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTATCCCCAAAGACAATGACATCGGCCCTCATGAAGCTGACGTTCCCACTCTCTCA	909
Qy	302	GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro	321
Db	910	GGCAAGCTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGAGGCTCACTCCAGCACCCCA	969
Qy	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341

Db	970	CTCTGGNATCATTTGGATGGGGCTTACGAGCAGCAATGGAGGGAGAGATGTCGTGACATACTG	1029
Qy	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaThrGln	351
Db	1030	CTCAGCGCGTCAGTCCAGGTCATTGACAGCACACGGTGCATGCAATGACAGCATTGGTACCCAG	1089
Qy	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1090	GGGGNAGTCACCCAGAGAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACCTGC	1149
Qy	382	GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle	401
Db	1150	CAGGGTGACAGTGGTGGGGCCCTGATGTACCAATCTGCACAGTGGCATGTGGTGGGCATC	1209
Qy	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1210	GTTAGCTGGGGGCTATGGCTCGGGGGCCCGAGCACCCAGGAGTATACACCAAGGCTCA	1269
Qy	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1270	GCCTATCTCAACTGGATCTACAAATGTCTGGAAGGCTGAGCTG	1311

RESULT 13

RES001.13
US-10-295-027-830
; Sequence 830, Application US/10295027
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patentin Ver. 2.1

Alignment Scores:	
Pred. No.:	0
Score:	2337.00
Length:	1314
Matches:	434

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 44 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-830 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuAryLysProArg 21
 DB 10 GATCTGACAGTATCAACCTCTGAACAGCCCTCGATGTCAAACCCCTCGCAAAACCCCGT 69
 QY 22 lleProMetGluThrPheArgLysValGlylleProLlellelleAlaLeuLeuSerLeu 41
 DB 70 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 129
 QY 42 AlaSerllellelleValValleuLysValleuLysValleuLysValleuLysValleu 61
 DB 130 GCGAGTATCATCTGTGTCTCATCAAGGTGATCTGGATATATATCTACTTCTCTC 189
 QY 62 CysGlyGlnProLeuHisPheleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
 DB 190 TCGGGCAGCCCTCTCCATCTATCCCGAGGAGCAGCTGTGTGACGGAGACTGGACTGT 249
 QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 DB 250 CCCCTGGGGAGGAGCAGACACTGTCTCAAGACTTCCCGAAGGCCCTGCACTGGCA 309
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 DB 310 GTCCGCCCTCTCCAAGACCGATCCACACTGCGAGGTCTGGACTCGGCCACAGGGAACCTG 369
 QY 122 PheSerAlaCysPheAspAsnPhetheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 DB 370 TTCCTCGCTGTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCGTGAGCAGATG 429
 QY 142 GlyTrpSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLeuAsp 161
 DB 430 GGCTACAGCAGCAACCCACTTTCAGACTGTGAGATTGGCCACAGACAGGATCTGGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 DB 490 GTTGTGAAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAATCAAGTGGGCCCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 DB 550 CTCTCAGGCTCCCTGGTCTCCCTGACTGTCTGCTGTGGGAAAGAGCTGGAAGACCCCC 609
 QY 202 ArgValValGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 DB 610 CGTGTGTGGTGGGGAGGAGCCCTCTGTGATTTCTTGGCCCTTGGCAGGTGAGCATCCAG 669
 QY 222 TyrAspLysGlnHisValCysGlySerlleLeuAspProHisTrpValLeuThrAla 241
 DB 670 TACGACAAACAGCAGCTGTGTGGAGGAGCATCTGGACCCCACTGGCTCTCACGGCA 729
 QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 DB 730 GCCCAGCTCTTCAGAAACATACCCGATGTGTCAACTGGAAGGTGCGGGCAGGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysllellellelleGluPheAsnPro 281
 DB 790 AAATGGGACGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCAACCCC 849
 QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 850 ATGTACCCAAAGACAATGACATCGCCCTCATGAAGCTGCAAGTTCCTCACTTCTCTCA 909
 QY 302 GlyThrValArgProLleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGCTCACTCCAGGCCACCCCA 969
 QY 322 LeuTrpIlelleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341

DB 970 CTCCTGGATCATTTGGATGGGCTTTACGACGACGAATGGAGGGAAGATGTCTGCATACTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
 DB 1030 CTGAGGCGCTCAGTCCAGGTCAITTCAGACGACACCGGTGCAATGACAGCATGCGTACCAG 1089
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlylleProGluGlyGlyValAspThrCys 381
 DB 1090 GGGGAAGTCCACCGAAGAAGATGATGTGTGCGAGGCATCCCGAAGGGGTGTGGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlylle 401
 DB 1150 CAGGTCACAGTGGTGGGCCCTTGATGTACATCTGACAGTGGCATGTGGTGGGATC 1209
 QY 402 ValSerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValTrpThrLysValSer 421
 DB 1210 GTTAGCTGGGCTATGCTGCGGGGCCCGAGCACCCCGAGAGTATACACCAAGGTCTCA 1269
 QY 422 AlaTrpLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1270 GCCTATCTCACTGATCTCAATGTCTGGAAGGCTGAGCTG 1311

RESULT 14

US-10-295-027-979
 ; Sequence 979, Application US/10295027
 ; GENERAL INFORMATION:
 ; APPLICANT: Afar, Daniel
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsberg, Wendy M.
 ; APPLICANT: Gish, Kurt C.
 ; APPLICANT: Glynn, Richard
 ; APPLICANT: Hevezi, Peter A.
 ; APPLICANT: Mack, David H.
 ; APPLICANT: Murray, Richard
 ; APPLICANT: Watson, Susan R.
 ; APPLICANT: Eos Biotechnology, Inc.
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
 ; FILE REFERENCE: 018501-012500US
 ; CURRENT FILING DATE: 2002-11-13
 ; CURRENT APPLICATION NUMBER: US/10/295,027
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 09/663,733
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/350,666
 ; PRIOR FILING DATE: 2001-11-13
 ; PRIOR APPLICATION NUMBER: US 60/335,394
 ; PRIOR FILING DATE: 2001-11-15
 ; PRIOR APPLICATION NUMBER: US 60/332,464
 ; PRIOR FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/334,393
 ; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340,376
 ; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347,211
 ; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/347,349
 ; PRIOR FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 60/355,250
 ; PRIOR FILING DATE: 2002-02-08
 ; PRIOR APPLICATION NUMBER: US 60/356,714
 ; PRIOR FILING DATE: 2002-02-13
 ; Remaining prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 1386
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 979
 ; LENGTH: 1314
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-295-027-979

Alignment Scores: 0 Length: 1314
 Pred. No.: 2337.00 Matches: 434
 Score:

Db 517 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGG 576
QY 122 PheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 577 TTCTCTGCTGTTTCGAACTTCACAGAACTCTCGCTGAGACAGCTGTAGGCAGATG 636
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLeuAsp 161
Db 637 GGCTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGAT 696
QY 162 ValValGluLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 697 GTTGTGTAATACAGAAACAGCCAGGAGCTTCGCATGCGAACTCAAGTGGGCCCTGT 756
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 757 CTCTCAGGCTCCCTGCTCTCCCTGCACTGCTTGGCTGTGGAGAGAGCTGAAGACCCCC 816
QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 817 CGTGTGCTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTGAGCATCCAG 876
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 877 TAGCACAAACAGCACGCTCTGTGGAGGAGCATCTGGACCCCACTGGGTCTCTCACGGCA 936
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 937 GCCCACTGCTTCAGGAAACATACCGATGTGTTCACTGGAGGTGGGGCAGGCTCAGAC 996
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 997 AAACCTGGCAGCTTCCCTGCTGCTGGCTGTGGCAAGATCATCATTTGAATTCACCC 1056
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1057 ATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTCCCACTCACTTTCTCA 1116
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1117 GGCACAGTCAGGCCCATCTGCTGCTGCTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1176
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1177 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGAGGGAGATGCTTGACATCTG 1236
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1237 CTGCAAGGCTCAGTCCAGTCAATTGACAGCACAGGTGCAATGACAGACGATGCGTACCAG 1296
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1297 GGGGAAGTCACCGAAGATGATGTGTCAGGCAATCCGGNAGGGGGTGTGGACACCTGC 1356
QY 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1357 CAGGTGACAGTGGTGGGCTTATGATGACAAATCTGACAGTGGCATGTGGTGGGCATC 1416
QY 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1417 GTTAGCTGGGCTATGGCTGGGGGGCCCGAGCACCCAGAGATATACCAAGGTCTCA 1476
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1518

Search completed: December 7, 2005, 00:56:55
Job time : 6365 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:45:59 ; Search time 576 Seconds
(without alignments)
1043.664 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVLRPKR.....VYTKVSAYLNWYNNWKAEL 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/PCTUS COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US056 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US073 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US074 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US075 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US076 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US077 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US078 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US079 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US080 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US081 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US082 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US083 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US084 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US085 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US086 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US087 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US088 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US089 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US090 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US091 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US092 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US093 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US094 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US095 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US096 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US097 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US098 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US099 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US100 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US101 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US102 COMB.pcp.*
- 33: /cgn2_6/ptodata/1/paa/US103 COMB.pcp.*
- 34: /cgn2_6/ptodata/1/paa/US104 COMB.pcp.*
- 35: /cgn2_6/ptodata/1/paa/US105 COMB.pcp.*
- 36: /cgn2_6/ptodata/1/paa/US106 COMB.pcp.*
- 37: /cgn2_6/ptodata/1/paa/US107 COMB.pcp.*
- 38: /cgn2_6/ptodata/1/paa/US108 COMB.pcp.*
- 39: /cgn2_6/ptodata/1/paa/US109 COMB.pcp.*
- 40: /cgn2_6/ptodata/1/paa/US110 COMB.pcp.*
- 41: /cgn2_6/ptodata/1/paa/US111 COMB.pcp.*
- 42: /cgn2_6/ptodata/1/paa/US112 COMB.pcp.*
- 43: /cgn2_6/ptodata/1/paa/US114 COMB.pcp.*

- 44: /cgn2_6/ptodata/1/paa/US600 COMB.pcp.*
- 45: /cgn2_6/ptodata/1/paa/US601 COMB.pcp.*
- 46: /cgn2_6/ptodata/1/paa/US602 COMB.pcp.*
- 47: /cgn2_6/ptodata/1/paa/US603 COMB.pcp.*
- 48: /cgn2_6/ptodata/1/paa/US604 COMB.pcp.*
- 49: /cgn2_6/ptodata/1/paa/US605 COMB.pcp.*
- 50: /cgn2_6/ptodata/1/paa/US606 COMB.pcp.*
- 51: /cgn2_6/ptodata/1/paa/US607 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2342	100.0	435	1	PCT-US01-18568-2	Sequence 2, Appli
2	2342	100.0	435	1	PCT-US01-18568A-2	Sequence 2, Appli
3	2342	100.0	435	1	PCT-US02-09671-1578	Sequence 1578, Ap
4	2342	100.0	435	1	PCT-US02-09671-1597	Sequence 1597, Ap
5	2342	100.0	435	30	US-10-030-688-2	Sequence 2, Appli
6	2342	100.0	435	34	US-10-473-127-1578	Sequence 1578, Ap
7	2342	100.0	435	34	US-10-473-127-1597	Sequence 1597, Ap
8	2342	100.0	435	38	US-10-803-530-2	Sequence 2, Appli
9	2342	100.0	461	1	PCT-US04-15258-3	Sequence 3, Appli
10	2342	100.0	461	1	PCT-US04-20741-7	Sequence 7, Appli
11	2338	99.8	435	1	PCT-US02-09671-1596	Sequence 1596, Ap
12	2338	99.8	435	26	US-09-659-151-6	Sequence 6, Appli
13	2338	99.8	435	31	US-10-180-719-6	Sequence 6, Appli
14	2338	99.8	435	34	US-10-473-127-1596	Sequence 6, Appli
15	2338	99.8	435	40	US-11-045-577-6	Sequence 6, Appli
16	2338	99.8	435	41	US-11-183-514-6	Sequence 6, Appli
17	2337	99.8	437	1	PCT-US02-09671-1581	Sequence 1581, Ap
18	2337	99.8	437	1	PCT-US02-09671-1586	Sequence 1586, Ap
19	2337	99.8	437	1	PCT-US02-09671-1601	Sequence 1601, Ap
20	2337	99.8	437	1	PCT-US02-09671-1602	Sequence 1602, Ap
21	2337	99.8	437	1	PCT-US02-19297-89	Sequence 89, Appli
22	2337	99.8	437	1	PCT-US04-21227-7	Sequence 7, Appli
23	2337	99.8	437	1	PCT-US04-38689-7	Sequence 7, Appli
24	2337	99.8	437	31	US-10-173-999-89	Sequence 89, Appli
25	2337	99.8	437	32	US-10-295-027-779	Sequence 779, App
26	2337	99.8	437	32	US-10-295-027-791	Sequence 791, App
27	2337	99.8	437	32	US-10-295-027-831	Sequence 831, App
28	2337	99.8	437	32	US-10-295-027-1196	Sequence 1196, Ap
29	2337	99.8	437	34	US-10-473-127-1581	Sequence 1581, Ap
30	2337	99.8	437	34	US-10-473-127-1586	Sequence 1586, Ap
31	2337	99.8	437	34	US-10-473-127-1601	Sequence 1601, Ap
32	2337	99.8	437	34	US-10-473-127-1602	Sequence 1602, Ap
33	2337	99.8	437	39	US-10-991-287-7	Sequence 7, Appli
34	2337	99.8	437	39	US-10-994-117-7	Sequence 7, Appli
35	2337	99.8	437	50	US-60-625-561-207	Sequence 207, App
36	2333	99.6	435	39	US-10-918-711-490	Sequence 490, App
37	2333	99.6	435	39	US-10-918-754-2200	Sequence 2200, Ap
38	2333	99.6	435	48	US-60-495-114-2200	Sequence 2200, Ap
39	2333	99.6	435	48	US-60-495-135-490	Sequence 490, App
40	2333	99.6	435	50	US-60-625-561-206	Sequence 206, App
41	2329	99.4	492	1	PCT-US02-09671-1579	Sequence 1579, Ap
42	2329	99.4	492	1	PCT-US02-09671-1598	Sequence 1598, Ap
43	2329	99.4	492	30	US-10-030-688-4	Sequence 4, Appli
44	2329	99.4	492	34	US-10-473-127-1579	Sequence 1579, Ap
45	2329	99.4	492	34	US-10-473-127-1598	Sequence 1598, Ap

ALIGNMENTS

RESULT 1
PCT-US01-18568-2
; Sequence 2, Application PC/TUS0118568
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L

; APPLICANT: Qi, Jian-shen
 ; APPLICANT: Andrade-Gordon, Patricia
 ; TITLE OF INVENTION: DNA encoding human serine protease D-G
 ; FILE REFERENCE: ORT-1273
 ; CURRENT APPLICATION NUMBER: PCT/US01/18568
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 435
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US01-18568-2

Query Match 100.0%; Score 2342; DB 1; Length 435;
 Best Local Similarity 100.0%; Pred. No. 3.2e-220;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVIDLKYTF	60
Db	1	MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVIDLKYTF	60
Qy	61	LCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN	120
Db	61	LCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN	120
Qy	121	WFSACFDNFTEALAEACRQMGYSKPTFRVVEIGPDQDLVVEITENSQELRMNNSGP	180
Db	121	WFSACFDNFTEALAEACRQMGYSKPTFRVVEIGPDQDLVVEITENSQELRMNNSGP	180
Qy	181	CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT	240
Db	181	CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLT	240
Qy	241	AAHCFRKHDTVFNWVKRAGSDKLGSPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF	300
Db	241	AAHCFRKHDTVFNWVKRAGSDKLGSPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPPLTF	300
Qy	301	SGTVRPICLPFFDEELTPATPLWIIIGWGTQKNGKMSDILLQASVQVIDSTRCNADDAY	360
Db	301	SGTVRPICLPFFDEELTPATPLWIIIGWGTQKNGKMSDILLQASVQVIDSTRCNADDAY	360
Qy	361	QGEVTEKMKCAGIPGEGVDTCCGDSGGPLMYQSDQHWVVGIVSGYGGGSPSTPGVYTKV	420
Db	361	QGEVTEKMKCAGIPGEGVDTCCGDSGGPLMYQSDQHWVVGIVSGYGGGSPSTPGVYTKV	420
Qy	421	SAYLNWIYNVWKAEL 435	
Db	421	SAYLNWIYNVWKAEL 435	

RESULT 3
 PCT-US02-09671-1578
 ; Sequence 1578, Application PC/TUS0209671
 ; GENERAL INFORMATION:
 ; APPLICANT: Zycos Inc.
 ; TITLE OF INVENTION: TRANSLATIONAL PROFILING
 ; FILE REFERENCE: 08191-026W01
 ; CURRENT APPLICATION NUMBER: PCT/US02/09671
 ; PRIOR FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: 60/279,495
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: 60/292,544
 ; PRIOR FILING DATE: 2001-05-21
 ; PRIOR APPLICATION NUMBER: 60/310,801
 ; PRIOR FILING DATE: 2001-08-08
 ; PRIOR APPLICATION NUMBER: 60/326,370
 ; PRIOR FILING DATE: 2001-10-01
 ; PRIOR APPLICATION NUMBER: 60/336,780
 ; PRIOR FILING DATE: 2001-12-04
 ; PRIOR APPLICATION NUMBER: 60/358,985
 ; PRIOR FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 2041
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1578
 ; LENGTH: 435
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US02-09671-1578

Query Match 100.0%; Score 2342; DB 1; Length 435;
 Best Local Similarity 100.0%; Pred. No. 3.2e-220;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVIDLKYTF	60
Db	1	MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVVLIKVIDLKYTF	60
Qy	61	LCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN	120

Db 61 LCGQPLHFIIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN 120
Qy 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMNNSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240
Qy 241 AAHCFRKHDTVFNWKRAGSKLGSFSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTVFNWKRAGSKLGSFSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPICIPLPFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Db 301 SGTVRPICIPLPFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAY 360
Qy 361 QGEVTEKMKCAGIPGGVDTCQDGGGGLMYQSDQHVHVGVISWGYGCGGSTPGVYTKV 420
Db 361 QGEVTEKMKCAGIPGGVDTCQDGGGGLMYQSDQHVHVGVISWGYGCGGSTPGVYTKV 420
Qy 421 SAYLNWIYVNWKAEL 435
Db 421 SAYLNWIYVNWKAEL 435

RESULT 4

PCT-US02-09671-1597

; Sequence 1597, Application PC/TUS0209671

; GENERAL INFORMATION:

; APPLICANT: Zycos Inc.

; TITLE OF INVENTION: TRANSLATIONAL PROFILING

; FILE REFERENCE: 08191-026W01

; CURRENT APPLICATION NUMBER: PCT/US02/09671

; CURRENT FILING DATE: 2002-03-28

; PRIOR APPLICATION NUMBER: 60/279,495

; PRIOR FILING DATE: 2001-03-28

; PRIOR APPLICATION NUMBER: 60/292,544

; PRIOR FILING DATE: 2001-05-21

; PRIOR APPLICATION NUMBER: 60/310,801

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: 60/326,370

; PRIOR FILING DATE: 2001-10-01

; PRIOR APPLICATION NUMBER: 60/336,780

; PRIOR FILING DATE: 2001-12-04

; PRIOR APPLICATION NUMBER: 60/358,985

; PRIOR FILING DATE: 2002-02-20

; NUMBER OF SEQ ID NOS: 2041

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1597

; LENGTH: 435

; TYPE: PRT

; ORGANISM: Homo sapiens

PCT-US02-09671-1597

Query Match 100.0%; Score 2342; DB 1; Length 435;

Best Local Similarity 100.0%; Pred. No. 3.2e-220;

Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVVKPLRPIMETFRKVGIPIIIALSLASIIIVVLIKVIDKYF 60
Db 1 MDPDSQPLNSLDVVKPLRPIMETFRKVGIPIIIALSLASIIIVVLIKVIDKYF 60
Qy 61 LCGQPLHFIIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN 120
Db 61 LCGQPLHFIIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAPAVRLSKDRSTLQVLDSATGN 120
Qy 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTALAEACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELRMNNSGP 180
Qy 181 CLSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLT 240

RESULT 6

```
US-10-473-127-1578
; Sequence 1578, Application US/10473127
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1578
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1578
```

```
Query Match      100.0%; Score 2342; DB 34; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVLIKVIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVLIKVIDKYFF 60
Qy 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSFPEGPAAVRLSKORSTLQVLD SATGN 120
Db 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSFPEGPAAVRLSKORSTLQVLD SATGN 120
Qy 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMN SSGP 180
Db 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMN SSGP 180
Qy 181 CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSI LDPHVVLT 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSI LDPHVVLT 240
Qy 241 AAHCFRKHTDVENKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Db 241 AAHCFRKHTDVENKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRC NADDAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRC NADDAY 360
Qy 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHVGVGIVSGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHVGVGIVSGYCGGPGSTPGVYTKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 7

```
US-10-473-127-1597
; Sequence 1597, Application US/10473127
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
```

```
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1597
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1597
```

```
Query Match      100.0%; Score 2342; DB 34; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVLIKVIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRPMPMETFRKVGPIIIALLSLASIIIVVLIKVIDKYFF 60
Qy 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSFPEGPAAVRLSKORSTLQVLD SATGN 120
Db 61 LCGQPLHFIPIKQLCDGELDCPLGEDEEHCVKSFPEGPAAVRLSKORSTLQVLD SATGN 120
Qy 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMN SSGP 180
Db 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMN SSGP 180
Qy 181 CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSI LDPHVVLT 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGGSI LDPHVVLT 240
Qy 241 AAHCFRKHTDVENKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Db 241 AAHCFRKHTDVENKVRAGSKLGSFPSPSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRC NADDAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRC NADDAY 360
Qy 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHVGVGIVSGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMKCAGIPEGGVDTCCQDGGPLMYQSDQHVGVGIVSGYCGGPGSTPGVYTKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
```

RESULT 8

```
US-10-803-530-2
; Sequence 2, Application US/10803530
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/10/803,530
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/09/607,745
; PRIOR FILING DATE: 2000-06-30
```

```
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.11
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-803-530-2

Query Match      100.0%; Score 2342; DB 38; Length 435;
Best Local Similarity 100.0%; Pred. No. 3.2e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60

Qy 61 LCGQPLHFIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120
Db 61 LCGQPLHFIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120

Qy 121 WFSACFNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGP 180
Db 121 WFSACFNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGP 180

Qy 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 240

Qy 241 AAHCFRKHTDVFNWVKRAGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 300
Db 241 AAHCFRKHTDVFNWVKRAGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 300

Qy 301 SGTVRPICLPFDELTATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPICLPFDELTATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360

Qy 361 QGEVTEKMKACGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMKACGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGCGGPGSTPGVYTKV 420

Qy 421 SAYLNWIYNNWKAEL 435
Db 421 SAYLNWIYNNWKAEL 435

RESULT 9
PCT-US04-15258-3
; Sequence 3, Application PC/TUS0415258
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr15 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: PCT/US04/15258
; PRIOR FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
PCT-US04-15258-3

Query Match      100.0%; Score 2342; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 3.4e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60

Qy 61 LCGQPLHFIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120
Db 61 LCGQPLHFIPRKQLCDGELDCPLGEDEBHCVKSPPEGPAVAVRLSKDRSTIQVLD SATGN 120

Qy 121 WFSACFNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGP 180
Db 121 WFSACFNFTEALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELRMRNSSGP 180

Qy 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 240

Qy 241 AAHCFRKHTDVFNWVKRAGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 300
Db 241 AAHCFRKHTDVFNWVKRAGSKLKTFRVVGGEASVDSWPMQVSIQYDKQHVCGGSILDPHWLT 300

Qy 301 SGTVRPICLPFDELTATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPICLPFDELTATPLWIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360

Qy 361 QGEVTEKMKACGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMKACGIPGEGVDTCQDSGGPLMYQSDQHVHVGVSVGCGGPGSTPGVYTKV 420

Qy 421 SAYLNWIYNNWKAEL 435
Db 421 SAYLNWIYNNWKAEL 435

PCT-US04-20741-7
; Sequence 7, Application PC/TUS0420741
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Papkoff, Jackie
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; APPLICANT: Kmet, Muriel
; APPLICANT: Tang, Jianwen
; TITLE OF INVENTION: Pro104 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0491
; CURRENT APPLICATION NUMBER: PCT/US04/20741
; CURRENT FILING DATE: 2004-07-06
; PRIOR FILING DATE: 2003-11-17
; PRIOR APPLICATION NUMBER: US 60/523,271
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US 60/485,346
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-20741-7

Query Match      100.0%; Score 2342; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 3.4e-220;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60
Db 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVLIKVLIDKYFF 60
```



```

QY 61 LCGPLHFI PRKQICDGLDCEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
DB 61 LCGPLHFI PRKQICDGLDCEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
DB 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
QY 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
DB 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
QY 241 AAHCFRKHDTDFVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHDTDFVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPCLPFPFDEELTPATPLIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
DB 301 SGTVRPCLPFPFDEELTPATPLIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QGEVTERKMCAGIPEGGVDTCCQSDGGLMYQSDQMHVVGVISWGYCGGSPSTPGVYTKV 420
DB 361 QGEVTERKMCAGIPEGGVDTCCQSDGGLMYQSDQMHVVGVISWGYCGGSPSTPGVYTKV 420
QY 421 SAYLNWIYNNWKAEL 435
DB 421 SAYLNWIYNNWKAEL 435

```

```

RESULT 11
PCT-US02-09671-1596
; Sequence 1596, Application PC/TUS0209671
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: PCT/US02/09671
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1596
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-09671-1596

```

```

Query Match 99.8%; Score 2338; DB 1; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVWLKILDKYF 60
DB 1 MDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPITIIALLSLASIIIVWLKILDKYF 60
QY 61 LCGPLHFI PRKQICDGLDCEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
DB 61 LCGPLHFI PRKQICDGLDCEDEHCHVKSPFPGPAVAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180

```

```

DB 121 WFSACFNFTEALAEACRQMGYSKPTFRAVEIGPDQDLVDVVEITENSQELMRNSSGP 180
QY 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
DB 181 CLSGLSVLSHCLACGKSLKTPRVVGGBEASVDSWPMQVSIQYDKQHVCGGSIILDPHWLT 240
QY 241 AAHCFRKHDTDFVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHDTDFVFNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPCLPFPFDEELTPATPLIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
DB 301 SGTVRPCLPFPFDEELTPATPLIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QGEVTERKMCAGIPEGGVDTCCQSDGGLMYQSDQMHVVGVISWGYCGGSPSTPGVYTKV 420
DB 361 QGEVTERKMCAGIPEGGVDTCCQSDGGLMYQSDQMHVVGVISWGYCGGSPSTPGVYTKV 420
QY 421 SAYLNWIYNNWKAEL 435
DB 421 SAYLNWIYNNWKAEL 435

RESULT 12
US-09-659-151-6
; Sequence 6, Application US/09659151
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/659,151
; FILING DATE: 11-Sep-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/008,271
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLN0013
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-659-151-6

```


Query Match 99.8%; Score 2338; DB 26; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 6 :

US-10-180-719-6

Query Match 99.8%; Score 2338; DB 31; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVPLKRPRIIPMETFRKVGIPPIIIALLSLASIIIVVLIKVIDKYF 60
DB 1 MDPDSQPLNSLDVPLKRPRIIPMETFRKVGIPPIIIALLSLASIIIVVLIKVIDKYF 60
QY 61 LCGQPLHFIIPRKQDCGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLOVLSATGN 120
DB 61 LCGQPLHFIIPRKQDCGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLOVLSATGN 120
QY 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNNSGP 180
DB 121 WFSACFDNTEALAEATACQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNNSGP 180
QY 181 CLSGSLVSLHCLACGSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
DB 181 CLSGSLVSLHCLACGSLKTPRVVGEEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
QY 241 AAHCFRKHDTVFNWVVRAGSKLGSPFSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHDTVFNWVVRAGSKLGSPFSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
DB 301 SGTVRPCLPFFDEELTPATPLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAY 360
QY 361 QGEVTEKMKCAGIPGEGVDTCQDGGPLMYOSDQHVHVGIVSWGVCGBPSTPGVYTKV 420
DB 361 QGEVTEKMKCAGIPGEGVDTCQDGGPLMYOSDQHVHVGIVSWGVCGBPSTPGVYTKV 420
QY 421 SAYLNWYNNVWKAEL 435
DB 421 SAYLNWYNNVWKAEL 435

RESULT 13

US-10-180-719-6
Sequence 6, Application US/10180719
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
Tang, Tom Y.
Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESS: Incyte Pharmaceuticals, Inc.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/180,719
FILING DATE: 25-Jun-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271
FILING DATE: 16-Jan-1998
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:

RESULT 14

US-10-473-127-1596
Sequence 1596, Application US/10473127
GENERAL INFORMATION:
APPLICANT: Zycos Inc.
TITLE OF INVENTION: TRANSLATIONAL PROFILING
FILE REFERENCE: 08191-026W01
CURRENT APPLICATION NUMBER: US/10/473,127
CURRENT FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: 60/279,495
PRIOR FILING DATE: 2001-03-28
PRIOR APPLICATION NUMBER: 60/292,544
PRIOR FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: 60/310,801
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: 60/326,370
PRIOR FILING DATE: 2001-10-01
PRIOR APPLICATION NUMBER: 60/336,780
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 60/358,985
PRIOR FILING DATE: 2002-02-20
NUMBER OF SEQ ID NOS: 2041

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1596
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1596

Query Match          99.8%; Score 2338; DB 34; Length 435;
Best Local Similarity 99.8%; Pred. No. 7.8e-220;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYIF 60
Db 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYIF 60
Qy 61 LCQGPLHFIPRKQDCGELDCPLGEDDEEHCVSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
Db 61 LCQGPLHFIPRKQDCGELDCPLGEDDEEHCVSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
Qy 241 AAHCFRKHDTDFNWKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTDFNWKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Qy 361 QGEVTEKMKCAGIPGGVDTCQDGGGLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
Db 361 QGEVTEKMKCAGIPGGVDTCQDGGGLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
Qy 421 SAYLNIYINVMKAEL 435
Db 421 SAYLNIYINVMKAEL 435
```

RESULT 15

US-11-045-577-6

; Sequence 6, Application US/11045577

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; Hillman, Jennifer L.

; Yue, Henry

; Guegler, Karl J.

; Corley, Neil C.

; Tang, Tom Y.

; Shah, Purvi

; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Dr.

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/11/045,577

; FILING DATE: 27-Jan-2005

; PRIOR APPLICATION DATA:

```
; APPLICATION NUMBER: US/09/659,151
; FILING DATE: 11-Sep-2000
; APPLICATION NUMBER: 09/008,271
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLANNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-11-045-577-6
```

Query Match

Best Local Similarity 99.8%; Score 2338; DB 40; Length 435;

Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYIF 60
Db 1 MDPDSQPLNSLDVVKLRPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYIF 60
Qy 61 LCQGPLHFIPRKQDCGELDCPLGEDDEEHCVSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
Db 61 LCQGPLHFIPRKQDCGELDCPLGEDDEEHCVSFPEGPAPAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGSILDPHWLT 240
Qy 241 AAHCFRKHDTDFNWKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Db 241 AAHCFRKHDTDFNWKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
Qy 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Db 301 SGTVRPILCPFFDEELTPATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTRCNAD DAY 360
Qy 361 QGEVTEKMKCAGIPGGVDTCQDGGGLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
Db 361 QGEVTEKMKCAGIPGGVDTCQDGGGLMYQSDQHVHVGIVSWGYGCGSPSTPGVYTKV 420
Qy 421 SAYLNIYINVMKAEL 435
Db 421 SAYLNIYINVMKAEL 435
```

Search completed: December 5, 2005, 14:04:27
Job time : 580 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 20:22:46 ; Search time 47 Seconds
(without alignments)
1186.911 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVKPLRP.....VYTKVSAYLNWYNWKAEL 435

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 52794 seqs, 64120477 residues

Total number of hits satisfying chosen parameters: 105588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlpl
-Q=/cgn2_1/USPTO_spool_p/US10803530/runat_05122005_083217_2215/app_query.fasta_1.583
-DB=Pending_Patents_NA_New -QFMT=fastap -SUFFIX=rnnp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530 @CEN_1_1.0 @runat_05122005_083217_2215 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLXQ=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:

- 1: /cgn2_6/ptodata/1/pna/PCT_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/US05_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US11_NEW_COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2333	99.6	2590	7	US-11-268-554-646 Sequence 646, App
2	2328	99.4	2627	7	US-11-268-554-645 Sequence 645, App
3	2297.5	98.1	2063	6	US-10-206-921A-329 Sequence 329, App
4	658	28.1	2412	6	US-10-206-921A-63 Sequence 63, Appl
5	574	24.5	2265	8	US-60-732-162-487 Sequence 487, App
6	478	20.4	1082	6	US-10-868-184C-2378 Sequence 2378, App
7	475	20.3	2123	6	US-10-206-921A-319 Sequence 319, App
8	475	20.3	5058	7	US-11-265-762-145 Sequence 145, App
9	463	19.8	3305	7	US-11-268-554-707 Sequence 707, App
10	460	19.6	2771	7	US-11-108-459-5 Sequence 5, Appli

11	443	18.9	2732	7	US-11-108-459-7	Sequence 7, Appli
12	356	15.2	1296	1	PCT-US05-39108-3	Sequence 3, Appli
13	347	14.8	4137	8	US-60-732-162-277	Sequence 277, App
14	323	13.8	4353	8	US-60-732-162-275	Sequence 275, App
15	299.5	12.8	2187	6	US-10-555-925-2	Sequence 2, Appli
16	299.5	12.8	2805	8	US-60-732-162-1603	Sequence 1603, Ap
17	299.5	12.8	2820	8	US-60-732-162-1599	Sequence 1599, Ap
18	298	12.7	1114	6	US-10-206-921A-271	Sequence 271, App
19	239.5	10.2	715	6	US-10-553-869-13	Sequence 53, Appl
20	233.5	10.0	920	7	US-11-134-141A-59	Sequence 169, App
21	226	9.6	2846	6	US-10-206-921A-169	Sequence 3874, Ap
22	202.5	8.6	400	7	US-11-267-914-3874	Sequence 12520, A
23	187.5	8.0	4493	6	US-10-868-184C-12520	Sequence 6222, Ap
24	175	7.5	7116	6	US-10-868-184C-6222	Sequence 7240, Ap
25	169.5	7.2	3374	6	US-10-868-184C-7240	Sequence 7238, Ap
26	169.5	7.2	7293	6	US-10-868-184C-7238	Sequence 7237, Ap
27	166.5	7.1	7280	6	US-10-868-184C-7237	Sequence 59, Appli
28	146	6.2	754	6	US-10-504-973-59	Sequence 9, Appli
29	143.5	6.1	5280	7	US-11-108-459-9	Sequence 1, Appli
30	138	5.9	357	4	US-08-452-190C-1	Sequence 28, Appl
31	135.5	5.8	241	7	US-11-267-348-28	Sequence 28, Appl
32	135.5	5.8	241	7	US-11-267-914-28	Sequence 2903, Ap
33	135.5	5.8	241	7	US-11-267-914-2903	Sequence 344, App
34	131.5	5.6	1461	6	US-10-868-184C-344	Sequence 7239, App
35	104.5	4.5	663	6	US-10-868-184C-7239	Sequence 792, App
36	100	4.3	541	6	US-10-868-184C-792	Sequence 10806, A
37	94	4.0	7286	6	US-10-868-184C-10806	Sequence 7244, Ap
38	91.5	3.9	760	6	US-10-868-184C-7244	Sequence 351, App
39	87.5	3.7	4407	6	US-10-206-921A-351	Sequence 6, Appli
40	86.5	3.7	3778	7	US-11-141-611-6	Sequence 457, App
41	86.5	3.7	6907	7	US-11-268-554-457	Sequence 456, App
42	85.5	3.7	6928	7	US-11-268-554-456	Sequence 8397, Ap
43	85.5	3.7	9261	6	US-10-868-184C-8397	Sequence 8, Appli
44	84.5	3.6	1446	7	US-11-179-363-8	Sequence 10, Appl
45	84.5	3.6	1446	7	US-11-179-363-10	

ALIGNMENTS

RESULT 1

US-11-268-554-646
; Sequence 646, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CLO01557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 646
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-268-554-646

Alignment Scores:
Pred. No.: 2.81e-237 Length: 2590
Score: 2333.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.62% Indels: 0
DB: 7 Gaps: 0

US-10-803-530-2 (1-435) x US-11-268-554-646 (1-2590)

Qy	1	MetAppProAspSerAspGlnProLeuAnSerLeuAspVallyAsProLeuArgLysPro	20
Db	226	ATGGATCTTCGACGATCAACTCTGACACCTCGATGTCACACCTCTGCAACCC	285
Qy	21	ArgileProMetGluThrPheArgLysValGlyLeProilelelelelelelelelele	40

Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro	201
Db	806	CTCTCAGGCTCCCTGGTCTCCCTGCATCTGTCTTGCCTGGGAAGACCCCTGAAGACCCCC	865
Qy	202	ArgValValGlyGlyGluGluAlaSerValAspSerTyrProTyrGlnValSerIleGln	221
Db	866	CGTGTGGTGGTGTGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTCAGCATCCAG	925
Qy	222	TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThrAla	241
Db	926	TAGCACAAACAGACAGCTGTGTGGAGGAGCATCTCTGCACCCCACTGGGTCTCACGGCA	985
Qy	242	AlaHisCysPheArgIysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	986	GCCCACTGCTTCAGGAACAATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCAGAC	1045
Qy	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro	281
Db	1046	AACTGGGACGCTTCCCATCTCCTGGCTGTGGCCNAGATCATCATATTGAATTCACCCC	1105
Qy	282	MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	1106	ATGTACCCCAAGACAATGATCGCCTCATGAAGTCGAGTTCCCACTCACTTTCTCA	1165
Qy	302	GlyThrValArgProIleCysLeuProPhePheAspGluLeuIleuThrProAlaThrPro	321
Db	1166	GGCACAGTCAGGCCCATCTGTCTGCCCTCTTTGATGAGGAGCTCACTCCAGCCACCCCA	1225
Qy	322	LeuTrpIleIleGlyTyrGlyPheThrLysIleAsnGlyGlyLysMetSerAspIleLeu	341
Db	1226	CTCTGGATCATTTGATGGGCTTTTACGAACGAAATGGAGGAAGATGTCTGACATACTG	1285
Qy	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln	361
Db	1286	CTGCAGGCGTCAGTCCAGTCAATGACAGCACACGGTGCAATGCAGACGATCGGTACCAG	1345
Qy	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys	381
Db	1346	GGGGAAGTCACCGAAGAAGATGATGTGTGCAGGCAATCCCGAAGGGGGTGTGGAACCTGC	1405
Qy	382	GlnClyAspSerGlyIleProLeuMetTyrClnSerAspGlnTyrHisValValGlyIle	401
Db	1406	CAGGGTGACAGTGTGTGGGCCCTCATGTATCCAAATCTGACACAGTGGCATGTGTGGGGATC	1465
Qy	402	ValSerTyrGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer	421
Db	1466	GTTAGTTGGGGCTATGGCTGGGGGGCCCGACGACCCCAAGAGTATACCAAGGTCTCA	1525
Qy	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu	435
Db	1526	GCCTATCTCACTGGATCTCAATGTCTTGAAGGCTGAGCTG	1567

RESULT 3

```

US-10-206-921A-329
; Sequence 329, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: 39780-3430RIC515
; CURRENT APPLICATION NUMBER: US/10/206, 921A
; PRIORITY FILING DATE: 2002-07-26
; PRIORITY APPLICATION NUMBER: US 10/052,586

```

```

; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 329
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-206-921A-329

Alignment Scores:
Pred. No.: 1.12e-233
Score: 2297.50
Percent Similarity: 98.85%
Best Local Similarity: 98.85%
Query Match: 98.10%
DB: 6
Matches: 429
Conservative: 0
Mismatches: 0
Indels: 5
Gaps: .1

```

US-10-803-530-2 (1-435) x US-10-206-921A-329 (1-2063)

Qy	2	AspProAspSerAspGlnProLeuAasnSerLeuAspValLysProLeuArgLysProArg	21
Db	219	GA TCCTGACAGTGATCAACCTCTGAA CAGCCTCGATGTCAAA CCCCCTGGCAAACCCCGT	278
Qy	22	IleProMetGluThrPheArgLysValcylLleProIlellelleAlaLeuLeuSerLeu	41
Db	279	ATCCCCATGGAGACCTTCAGAAAGTGGGGATCCCATCATAGCACTACTGAGCCTG	338
Qy	42	AlaSerllellelleValValLeuLleLysVallleLeuAspLysTyrTyrPheLeu	61
Db	339	GCGAGTATCATATTGTGGTTGTCTCATCAAGGTGATTCGGATAAACTACTACTTCCTC	398
Qy	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys	81
Db	399	TGCGGGGAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGA	458
Qy	82	ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla	101
Db	459	CCCTTGGGGAGGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGGCGCTGCAGTGGCA	518
Qy	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
Db	519	GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTCTGGACTCGGCCACAGGAACTGG	578
Qy	122	PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet	141
Db	579	TTCTCTGCCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATG	638
Qy	142	GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp	161
Db	639	GGCTACAGC-----AGAGCTGTGGAGATTGGCCACAGACAGGATCTGGAT	683
Qy	162	ValValGluIleThrGluAasnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys	181
Db	684	GT TTGTGAAATCACAGAAAA CAGCCAGGAGCTTCGCATGCGGAAC TCAAGTGGGCCCTGT	743
Qy	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	744	CTCTCAGGCTCCCTGGTCTCCCTGCAC TGTCTTGGCTGTGGGAAGAGCCTGAAGACCCC	803
Qy	202	ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
Db	804	CGTGTGGTGGTGGGAGGAGGCCCTCTGTGGATTCTTGGCCTTGGCAGGTCAAGATCCAG	863
Qy	222	TyrAspLysGlnHisValCysGlyGlySerlleLeuAspProHisTrpValLeuThrAla	241
Db	864	TACGACAAACAGCAGCTCTGTGGAGGGAGCATCTCTGGACCCCACTGGGTCTCTACGGCA	923
Qy	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp	261
Db	924	GCCCACTGCTTTCAGGAAACATACCGATGTGTTCAACTGGAAGTGGCGAGGCTCAGAC	983


```
US-10-868-184C-2378
; Sequence 319, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; CURRENT FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2378
; LENGTH: 1082
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-868-184C-2378

Alignment Scores:
Pred. No.: 4,378-42 Length: 1082
Score: 478.00 Matches: 101
Percent Similarity: 55.81% Conservative: 48
Best Local Similarity: 37.83% Mismatches: 86
Query Match: 20.41% Indels: 32
DB: 6 Gaps: 8

US-10-803-530-2 (1-435) x US-10-868-184C-2378 (1-1082)
Qy 178 SerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSer 197
Db 97 TCAGGACCA-----TGGCGCGGACGG 117
Qy 198 LeuLysThrProArgValValGlyGluGluAlaSerValAspSerTrpProTrpGln 217
Db 118 GTCAATCAGTCGCGCATCGTGGTGGAGAGCGCGAACTCGGCGGTGGCGTGGCAG 177
Qy 218 ValSerIleGlnTyAspLysGlnHisValCysGlySerIleLeuAspProHisTrp 237
Db 178 GGGAGCTGCGCCCTGGGATCCACGATGCGGAGTGAGCGCTGCTACGCCACCGCTGG 237
Qy 238 ValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn-----Tip 254
Db 238 GCACTCAGCGGGCGCACTGCTTTGAAACCTATAGTACCTTAGTGATCCCTCGGGTGG 297
Qy 255 LysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLys--- 273
Db 298 ATGCTCCAGTTGGC-----CAGCTGACTTCCATGCCATCTCTGAGCGCTGCAGGCC 351
Qy 274 -----IleIleIleIleGluPheAsnProMetTyTrpProLysAspAsn 287
Db 352 TACTACACCGGTACTTCGTATCGAATATTAICTGAGCCCTCGCTACCTGGGGAATCA 411
Qy 288 -----AspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
Db 97 TCAGGACCA-----TGGCGCGGACGG 117

US-10-206-921A-319
; Sequence 319, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 319
; LENGTH: 2103
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-206-921A-319

Alignment Scores:
Pred. No.: 2,398-41 Length: 2103
Score: 475.00 Matches: 97
Percent Similarity: 56.33% Conservative: 41
Best Local Similarity: 39.59% Mismatches: 97
Query Match: 20.28% Indels: 10
DB: 6 Gaps: 5

US-10-803-530-2 (1-435) x US-10-206-921A-319 (1-2103)
```



```
QY 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 557 CATTCG-----TGCGGAACACGAAAGTAACCTCTAGGTTCAGAGTCTCAGGATCGTT 610
QY 205 GlyGlyGluGluAlaSerValAspSerTrpProTTPGlnValSerIleGlnTyrAspLys 224
Db 611 GGTGGACAGAGTAAAGAGGGTGAATGGCCCTGGCAGGTAGCTGCAGTGGGATGGG 670
QY 225 GlnHisValCysGlyCysIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 671 AGTCATCGCTGGGACCAACCTTAATAATGCCACATGGCTGTGAGTGTCTGCTCACTGT 730
QY 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264
Db 731 TTTCAACATATAAGAACCTCCAGATGAGCTGCTCTCTGGAGTAAACAATAAAACCT 790
QY 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 791 TCGAAATGAACGGGGTCTCGGAGATTAATGTCCATGAATAAATACAAACCCATCA 850
QY 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 851 CATGACTATGATATTCTCTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 910
QY 305 ArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIle 324
Db 911 CATAGAGTTTGTCTCCCTGATGCATCTCTAGTATTCAACCAAGGTGATGTGTTGTG 970
QY 325 IleGlyTrpGlyPheThrLysGlnAnGlyGlyLysMetSerAspIleLeuGlnAla 344
Db 971 ACAGAGTTTGGAGCTGAAA---AATGATGTTACAGTCAAAATCATCTTCGACAGCA 1027
QY 345 SerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluVal 364
Db 1028 CAGGTGACTCTCATAGACCTACAACCTTCAATGAACCTCAAGCTTACAATCAGGCCATA 1087
QY 365 ThrGluLysMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1088 ACTCTAGAATGTTATGTCTGCTGGCTCTTAGAAGGAAAAACAGATGCATGCCAGGGTGAC 1147
QY 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1148 TCTGGAGGACCACTGTTAGTTACAGATGCTAGAGATATCTGGTACCTCTCTGGAATAGTG 1207
QY 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1208 ACTGGGGAGATCAATGTGCGAAACCCACAAAGCCCTGGTGTATTACTAGATTACGGCC 1267
QY 423 TyrLeuAsnTrpIle 427
Db 1268 TTGCGGGAGCTGGATT 1282

RESULT 8
US-11-265-762-145
; Sequence 145, Application US/11265762
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/11/265,762
; CURRENT FILING DATE: 2005-11-02
; PRIOR APPLICATION NUMBER: US/09/796,753
; PRIOR FILING DATE: 2004-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
```

```
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 145
; LENGTH: 5058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-265-762-145
```

```
Alignment Scores:
Pred. No.: 8,58e-41 Length: 5058
Score: 475.00 Matches: 97
Percent Similarity: 56.33% Conservative: 41
Best Local Similarity: 39.59% Mismatches: 97
Query Match: 20.28% Indels: 10
DB: 7 Gaps: 5
```

US-10-803-530-2 (1-435) x US-11-265-762-145 (1-5058)

```
QY 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 556 CATTCG-----TGCGGAACACGAAAGTAACCTCTAGGTTCAGAGTCTCAGGATCGTT 619
QY 205 GlyGlyGluGluAlaSerValAspSerTrpProTTPGlnValSerIleGlnTyrAspLys 224
Db 620 GGTGGACAGAGTAAAGAGGGTGAATGGCCCTGGCAGGTAGCTGCAGTGGGATGGG 679
QY 225 GlnHisValCysGlyCysIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 680 AGTCATCGCTGGGAGCAACCTTAATAATGCCACATGGCTGTGAGTGTCTGCTCACTGT 739
QY 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264
Db 740 TTTCAACATATAAGAACCTCCAGATGAGCTGCTCTCTGGAGTAAACAATAAAACCT 799
QY 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 800 TCGAAATGAACGGGGTCTCGGAGATAAATGTCCATGAATAAATACAAACCCATCA 859
QY 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 860 CATGACTATGATATTCTCTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 919
QY 305 ArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIle 324
Db 920 CATAGAGTTTGTCTCCCTGATGCATCTATAGATTCAACCAAGGTGATGTGTTGTG 979
QY 325 IleGlyTrpGlyPheThrLysGlnAnGlyGlyLysMetSerAspIleLeuGlnAla 344
Db 980 ACAGAGTTTGGAGCACTGAAA---AATGATGTTTACAGTCAAAATCATCTTCGACAAGCA 1036
QY 345 SerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluVal 364
Db 1037 CAGGTGACTCTCATAGACCTACAACCTTGAATGAACCTCAAGCTTACAATCAGGCCATA 1096
QY 365 ThrGluLysMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1097 ACTCTAGAATGTTATGTCTGCTGGCTCTTAGAAGGAAAAACAGATGCATGCCAGGGTGAC 1156
QY 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1157 TCTGGAGGACCACTGTTAGTTACAGATGCTAGAGATATCTGTTACTGGAATAGTG 1216
QY 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1217 AGCTGGGAGATGAATGTGCGAAACCCACAAAGCCCTGGTGTATTACTAGATTACGGCC 1276
```

```

Qy 423 TyrLeuAsnTrpIle 427
Db 1277 TTGGGAGTGGATT 1291

RESULT 9
US-11-268-554-707
; Sequence 707, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 707
; LENGTH: 3305
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-268-554-707

Alignment Scores:
Pred. No.: 8 52e-40 Length: 3305
Score: 463.00 Matches: 123
Percent Similarity: 48.21% Conservative: 65
Best Local Similarity: 31.54% Mismatches: 151
Query Match: 19.77% Indels: 52
DB: 7 Gaps: 15

US-10-803-530-2 (1-435) x US-11-268-554-707 (1-3305)

Qy 58 TyrTyrPheLeuCysGly-----GlnProLeuHisPheIleProArgLysGln 73
Db 1667 CACCAGTTCCAGTCGAAGAACAAAGTTCTGCAAGCCCTCTTCTGGTGC----- 1714

Qy 74 LeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysValLysSer 93
Db 1715 ----TGGCAGGTGTGACGACTGCGGAGACACACGACGACGAGCGGGTGC----- 1762

Qy 94 PheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnVal 113
Db 1763 -----AGTTGTCGCGCCACGACTTCTAGGTGTTCCAAATGGAGTGC-----CTC 1807

Qy 114 LeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeu 133
Db 1808 TCGAAAAGCCAGCAGTGAATGGGAAGGACGACTGTGGGACGGTCCGACGAGGCC--- 1864

Qy 134 AlaGluThrAlaCysArgGlnMetGlyTyr-----SerSerLysProThrPheArgAla 151
Db 1865 -----TCTGCCCCAAGGTGAACGTGCTGCTACTGTGTACCAACACACACTACCGTGC 1915

Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 1916 -----CTCAATGGCTCTGCTTGACCAAGGCAACCCCTGAG 1951

Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySer-LeuValSerLeuHisCy 191
Db 1952 TGT-----GACGGGAAGGAGGACTGTAGCGCGCTCAGATGAGAAGGACTGCGAC 2002

Qy 191 sLeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVa 211
Db 2003 TGTGGGTGCGGTCTATTACG-AGACAGGCTGTTGTTGGGGGACGAGATCGGATGA 2061

Qy 211 LAspSerTrpProTrpGlnValSerIleGln---TyrAspLysGlnHisValCysGlyG1 230
Db 2062 GGGCAGTGGCCCTGGCAGGTGAAGCTGCTGCTGGCGCAGGCGCCACATCTGCGGTGC 2121

Qy 230 ySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe----- 245
Db 2122 TTCCCTCATCTCCCACTGGCTGGTCTCTGCGGCACACTGCTACATGATGACAGAGG 2181

```

```

Qy 246 -ArgLysHisThrAspValPheAsnTrpLysValAlaGly-----SerAspLysLe 263
Db 2182 ATTACAGTACTCAGACCCACAGTCAGCGCTTCTCTGGCTTGCACGACGAGGCCA 2241

Qy 263 uGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTy 283
Db 2242 GCGCAGCGCCCTCGGGTGCAGGAGCGAGCTCAAGCGCATCATCTCCACCCCTTCTT 2301

Qy 283 rProLys-----AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPh 300
Db 2302 CAATGACTTACCTTCGACTATGACATCGCGTCTGGAGCTGGAGAAACCGGACAGTA 2361

Qy 300 eSerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaTh 320
Db 2362 CAGCTCCATGTTGGTGGCGCCATCTGCTGCGGACGCTCCCATCTCTCTCCCTGCGGCAA 2421

Qy 320 rProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAsp11 340
Db 2422 GGCCATCTGGGTACGCGGCTGGGGACACAC---CAGTATGGAGGCACTGGCGCGTGTAT 2478

Qy 340 eLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTy 360
Db 2479 CCTGAAAAGGTTGAGATCGCGCTCATCAACACGACACCTGC-----CAGAACCTCCT 2532

Qy 360 rGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspTh 380
Db 2533 GCGCAGCAGATCAGCGCGCATGATGCTGCTGGGCTTCTCAGCGCGCGCTGGACTC 2592

Qy 380 rCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisVal---- 398
Db 2593 CTGCCAGGTGATTCGGGGGACCCCTGTCCAGCGTGGAGCGGATGGGGGATCTTCCA 2652

Qy 399 -ValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThr 418
Db 2653 GCGCGGTGTGTGAGTGGGAGACGCTGCGCTCAGAGGAACAACGACGCGGTGTACAC 2712

Qy 418 rLysValSerAlaTrpLeuAsnTrpIle 427
Db 2713 AAGGCTCCCTCTGTTTCGGGACTGGATC 2740

RESULT 10
US-11-108-459-5
; Sequence 5, Application US/11108459
; GENERAL INFORMATION:
; APPLICANT: Zaas, Aimee
; APPLICANT: Schwartz, David A.
; APPLICANT: Peltz, Gary
; TITLE OF INVENTION: Polymorphic Plasminogen Genes and Uses Thereof
; FILE REFERENCE: DUKE-03785
; CURRENT APPLICATION NUMBER: US/11/108,459
; CURRENT FILING DATE: 2005-04-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5
; LENGTH: 2771
; TYPE: DNA
; ORGANISM: Mus musculus
US-11-108-459-5

Alignment Scores:
Pred. No.: 1.37e-39 Length: 2771
Score: 460.00 Matches: 121
Percent Similarity: 48.66% Conservative: 61
Best Local Similarity: 32.35% Mismatches: 127
Query Match: 19.64% Indels: 66
DB: 7 Gaps: 12

US-10-803-530-2 (1-435) x US-11-108-459-5 (1-2771)

Qy 76 AspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysValLysSerPhePro 95
Db 1486 GACTCTGAGACAGACTGCTATGATGGGAATGGCAAGACTAT----- 1527

```



```
Db 1986 TCTCGAACCCGATGTTTCAGAAATAGAAAGTGTCTAGGCTGTTCC----- 2028
Qy 281 oMetTyrProLysAspAsnAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 2029 -TTGGAGCCACACGACAAAGATATTGCTTGTCTAAAGCTTAAGCAGTCTCGCGTCATCAC 2087
Qy 301 rGlyThrValArgProLysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 2088 TGACAAAGTAATCCCAAGCTTGTCTGCATCCCCCAATATTATGCTGCTGACCGGACGA 2147
Qy 321 oLeuTrpIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 2148 ATGTTTCATCAGTGGTGGGGAACCAAGTACTTTTGGG-----GCTGGCCTTCT 2201
Qy 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrG1 361
Db 2202 CAAGGAAGCCGACTCCCTGTGATTGAGATAAAGTGTGCAATCGCTATGAGTTTCTGAA 2261
Qy 361 nGlyGluValThrGluLysMetMetCysAlaGlyLleProGluGlyValAspThrCy 381
Db 2262 TGAAGAGTCCCAATCCACCGAACTCTGTCTGGCATTTGGCGGAGGCACTGACAGTTG 2321
Qy 381 sGlnGlyAspSerGlyGlyProLeuMet--TyrGlnSerAspGlnTrpHisValValG1 400
Db 2322 CCAGGCTGACAGTGGAGTCTCTGTGTTGCTTCGAGAGGACAATACATTTTACAAGG 2381
Qy 400 yIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysVa 420
Db 2382 AGTCACATCTTGGGGTCTTGGCTGTGCACGCCCAATAAGCTGTGTCTATGTCGTTG 2441
Qy 420 lSerAlaTyrLeuAsnTrpIleTyrAsnValTrpLys 432
Db 2442 TTCAAGTTTGTACTTGATTGAGGAGTGATGAGA 2478
```

RESULT 12

```
PCT-US05-39108-3
; Sequence 3, Application PC/TUS0539108
; GENERAL INFORMATION:
; APPLICANT: Abraham, Edward
; TITLE OF INVENTION: Antibodies That Bind Urokinase-Type Plasminogen Activator and
; FILE OF INVENTION: Epitopes Therefor
; FILE REFERENCE: 322690
; CURRENT APPLICATION NUMBER: PCT/US05/39108
; PRIOR FILING DATE: 2005-11-07
; PRIOR APPLICATION NUMBER: 60/623,188
; PRIOR FILING DATE: 2004-10-29
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; TYPE: DNA
; LENGTH: 1296
; ORGANISM: Homo sapiens
PCT-US05-39108-3
```

```
Alignment Scores:
Pred. No.: 4.25e-29 Length: 1296
Score: 356.00 Matches: 91
Percent Similarity: 49.25% Conservative: 40
Best Local Similarity: 34.21% Mismatches: 103
Query Match: 15.20% Indels: 32
DB: 1 Gaps: 9
```

```
US-10-803-530-2 (1-435) x PCT-US05-39108-3 (1-1296)
```

```
Qy 194 CysGlyLysSerLeuLysThrProArg-----ValValGlyGlyGluGluAlaSerVal 211
Db 502 TGTGGCCAAAGACTCTGAGGCCCGCTTTAAGATTATTGGGGAGAAATTCACCACCATC 561
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyrAspLysGln----- 225
Db 562 GAGAACCAAGCCCTGTTTGGGCCATC--TACAGAGGACCGGGGGGCTCTGTCAACC 618
```

```
Qy 226 HisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHiCysPhe 245
Db 619 TACGTGTGTGGAGGAGCCTCATAGCCCTTGTGGGTGATCAGCGGCACACACTGCTTC 678
Qy 246 ArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySer 265
Db 679 ATTGATTACCCAAAGAGGAGGACTACATCGTCTACTCGGTGCTCAAGGCTTAACCTCC 738
Qy 266 PhePro-----SerLeuAlaValAlaLysIleIleLeuGluPheAsnPro 281
Db 739 AACACGCAAGGGGAGATGAAGTTTGGGTGAAACCTCATCTCA-----CACAA 789
Qy 282 MetTyrProLysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPhe 296
Db 790 GACTACAGCGCTGACAGCTTGTCTACCAACAGCATTTGCTTGTGAAGATCCGTTCC 849
Qy 297 -----ProLeuThrPheSerGlyThrValArgProIleCysLeu 309
Db 850 AAGGAGGGCAGGTGTGGCGAGCCA-----TCCGGACTATACAGACCATCTGCTG 900
Qy 310 ProphePheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPhe 329
Db 901 CCTCGATGATTAACGATCCCCAGTTTGGCACCAAGCTGTGAGATCACTGGCTTTGAAAA 960
Qy 330 ThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIle 349
Db 961 GAGAAATCTACCGACTATCTCTATCCGGAGCAGCTGAAAATGACTGTTTGAAGCTGATT 1020
Qy 350 AspSerThrArgCysAsnAlaAspAlaTyrGlnGlyValThrGluLysMetMet 369
Db 1021 TCCACCGGGAGTGTGAGCAGCCCCACTACTACGGCTCTGAAGTACCACCAAAATGCTG 1080
Qy 370 CysAlaGlyLleProGluGlyValAspThrCysGlnGlyAspSerGlyGlyProLeu 389
Db 1081 TGTGCTGTGACCCACAGTGGAACAGATTCTCCAGGGAGACTCAGGGGAGCCCTC 1140
Qy 390 MetTyrGln---SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCys 408
Db 1141 GTCTGTTCCTCCAAAGCCGCATGACTTTGACTGGAATTGTGAGCTGGGGCGGTGATGT 1200
Qy 409 GlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyr 428
Db 1201 GCCTCGAAGGACAAAGCCAGCGCTCTACAGAGAGTCTACACTTCTTACCTGGATCCGC 1260
Qy 429 AsnValTrpLysAlaGlu 434
Db 1261 AGTCACACCAAGGAGAG 1278
```

RESULT 13

```
US-60-732-162-277
; Sequence 277, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; FILE OF INVENTION: Genemap of the Human Genes Associated with Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60/732,162
; CURRENT FILING DATE: 2005-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 277
; LENGTH: 4137
```

```

; TYPE: DNA
; ORGANISM: Homosapiens
US-60-732-162-277

Alignment Scores:
Pred. No.: 2.06e-27 Length: 4137
Score: 347.00 Matches: 122
Percent Similarity: 43.69% Conservative: 58
Best Local Similarity: 29.61% Mismatches: 138
Query Match: 14.82% Indels: 94
DB: 8 Gaps: 23

US-10-803-530-2 (1-435) x US-60-732-162-277 (1-4137)
Qy 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 1272 CCTCCCAAGCCCAAGTATTTCTTCAAGAC-----CAAGTCTGCTGACGCTGT 1319
Qy 118 ---ThrGlyAsnTrpPheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThr 136
Db 1320 GACACAGC-----TACAAAGTGCTGAAGGATAATGTGGAGATGACACATTCACAGT 1373
Qy 137 AlaCysArgGlnMetGly---TyrSerLys---ProThrPheArgAlaValGluLeu 154
Db 1374 GAGTCTCTGAAGGATGGCAGCGGAGTAAACAGATTCCACCTGTAAATTTGTAGACTGT 1433
Qy 155 GlyProAspGlnAspLeuAsp-----ProCysLeuSerGly 184
Db 1434 AGAGCCCGGAGGAGCTGGACACACGCGGTGATCACCTTCTTCAAGGAACAACCTCAC 1493
Qy 162 -----ValValGluLeuThrGluAsnSerGlnGlu-----LeuArgMetArgAsn 176
Db 1494 ACAATCAAGTCTGAGATCAATACTCTCTGTCAGGAGCCCTATTACAGATGCTCAACAT 1553
Qy 177 SerSerGly-----ProCysLeuSerGly 184
Db 1554 AACACAGGTATATATACCTGTTCTGCCCAAGGAGTCTGGATGAATAAAGTATTGGGAGA 1613
Qy 185 SerLeuValSerLeuHisCysLeu---AlaCysGlyLysSerLeuLysThr----- 200
Db 1614 AGCTTACCAC-----TGCCTTCCAGAGTGTGGTCAAGCTCCCGCTCCCTGCCAAGC 1667
Qy 201 -----ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnVal 218
Db 1668 CTGGTCAAGAGGATCATTCGGGGCGGAAATGTGAGCGCTCTTCCCGTGGCAGGCC 1727
Qy 219 SerIleGlnTyrAsp-----LysGlnHisValCysGlyGlySer 231
Db 1728 CTGATAGTGTGGAGACACTTCGAGAGTCCCAATGACAGAGTGGTTTGGAGTGGGCC 1787
Qy 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArg----- 246
Db 1788 CTGCTCTCTGCTGCTGGATCTCTCACAGCAGCTCATGTGCTGGCTCCCGAGGTGAGAC 1847
Qy 247 -----LysHisThrAspValPheAsnTrpLysValArgAla 258
Db 1848 ACCACGGTGATACCACTCTCCAAAGGAGCATGTCCACCGTCTAC---CTGGGCTTGCATGAT 1904
Qy 259 GlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlu 278
Db 1905 GTGCGAGACAAATCGGGGCGAGTCAACAGCTCAGCTGCGCGGAGTGTGCTCCACCCAGAC 1964
Qy 279 PheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLysGlnPheProLeu 298
Db 1965 TTCAAC---ATCCAAACATCAACACCATATAGCTCTGTGTGAGCTGCGAGGAGCTGTG 2021
Qy 299 ThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrPro 318
Db 2022 CCCTGGGACCCACAGTTATGCTGTCTGCTGCCAAGGCTTGAGCGCTGAA---GGCCCG 2078
Qy 319 AlaThrProLeu-----TrpIleIleGlyTrpGlyPheThrLysGlnAsn----- 333
Db 2079 GCCCCCCACATGCTGGGCGCTGGTGGCGGCTGGGGCATCTCCAATCCCAATGTGACAGTG 2138

RESULT 14
US-60-732-162-275
; Sequence 275, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60732,162
; CURRENT FILING DATE: 2003-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 275
; LENGTH: 4353
; TYPE: DNA
; ORGANISM: Homosapiens
US-60-732-162-275

Alignment Scores:
Pred. No.: 7.55e-25 Length: 4353
Score: 323.00 Matches: 109
Percent Similarity: 42.75% Conservative: 65
Best Local Similarity: 26.78% Mismatches: 133
Query Match: 13.79% Indels: 100
DB: 8 Gaps: 20

US-10-803-530-2 (1-435) x US-60-732-162-275 (1-4353)
Qy 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 1272 CCTCCCAAGCCCAAGTATTTCTTCAAGAC-----CAAGTCTGCTGACGCTGT 1319
Qy 118 ---ThrGlyAsnTrpPheSerAlaCysPheAspAsnThrGluAlaLeuAlaGluThr 136
Db 1320 GACACAGC-----TACAAAGTGCTGAAGGATAATGTGGAGATGACACATTCACAGT 1373
Qy 137 AlaCysArgGlnMetGly---TyrSerLys---ProThrPheArgAlaValGluLeu 154
Db 1374 GAGTCTCTGAAGGATGGCAGCGGAGTAAACAGATTCCACCTGTAAATTTGTAGACTGT 1433
Qy 155 GlyProAspGlnAspLeuAsp-----ProCysLeuSerGly 184
Db 1434 AGAGCCCGGAGGAGCTGGACACACGCGGTGATCACCTTCTTCAAGGAACAACCTCAC 1493
Qy 162 -----ValValGluLeuThrGluAsnSerGlnGlu-----LeuArgMetArgAsn 176
Db 1494 ACAATCAAGTCTGAGATCAATACTCTCTGTCAGGAGCCCTATTACAGATGCTCAACAT 1553
Qy 177 SerSerGly-----ProCysLeuSerGly 184
Db 1554 AACACAGGTATATATACCTGTTCTGCCCAAGGAGTCTGGATGAATAAAGTATTGGGAGA 1613
Qy 185 SerLeuValSerLeuHisCysLeu---AlaCysGlyLysSerLeuLysThr----- 200
Db 1614 AGCTTACCAC-----TGCCTTCCAGAGTGTGGTCAAGCTCCCGCTCCCTGCCAAGC 1667
Qy 201 -----ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnVal 218
Db 1668 CTGGTCAAGAGGATCATTCGGGGCGGAAATGTGAGCGCTCTTCCCGTGGCAGGCC 1727
Qy 219 SerIleGlnTyrAsp-----LysGlnHisValCysGlyGlySer 231
Db 1728 CTGATAGTGTGGAGACACTTCGAGAGTCCCAATGACAGAGTGGTTTGGAGTGGGCC 1787
Qy 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArg----- 246
Db 1788 CTGCTCTCTGCTGCTGGATCTCTCACAGCAGCTCATGTGCTGGCTCCCGAGGTGAGAC 1847
Qy 247 -----LysHisThrAspValPheAsnTrpLysValArgAla 258
Db 1848 ACCACGGTGATACCACTCTCCAAAGGAGCATGTCCACCGTCTAC---CTGGGCTTGCATGAT 1904
Qy 259 GlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlu 278
Db 1905 GTGCGAGACAAATCGGGGCGAGTCAACAGCTCAGCTGCGCGGAGTGTGCTCCACCCAGAC 1964
Qy 279 PheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLysGlnPheProLeu 298
Db 1965 TTCAAC---ATCCAAACATCAACACCATATAGCTCTGTGTGAGCTGCGAGGAGCTGTG 2021
Qy 299 ThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrPro 318
Db 2022 CCCTGGGACCCACAGTTATGCTGTCTGCTGCCAAGGCTTGAGCGCTGAA---GGCCCG 2078
Qy 319 AlaThrProLeu-----TrpIleIleGlyTrpGlyPheThrLysGlnAsn----- 333
Db 2079 GCCCCCCACATGCTGGGCGCTGGTGGCGGCTGGGGCATCTCCAATCCCAATGTGACAGTG 2138
```



```
Db 1819 CCTGAAAAGACCAGTTGCAGTGTATTATGGCTGGGGCTACACT-----GGATTG 1866
Qy 337 MetSer-----AspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCys 354
   ::::| | | | | ::::| | | | | ::::| | | | | ::::| | | | |
Db 1867 ATCACTATGATGGCTATTACGAGTGGCACATCTCTATATAATGGGAAATGAGAAATGC 1926
Qy 355 AsnAlaAspAspAlaTyrGlnGluValThr-----GluLysMetMetCysAlaGly 372
   ::::| | | | | ::::| | | | | ::::| | | | | ::::| | | | |
Db 1927 AGC-----CAGCATCATCGAGGGAAGGTGACTCTGAATGAGTCTGAAATATGTCTGGG 1980
Qy 373 IleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGln 392
   ::::| | | | | ::::| | | | | ::::| | | | | ::::| | | | |
Db 1981 GCTGAAAAGATTGGATCAGGACCATGTGAGGGGATATGTTGGCCCACTTGTTTGTGAG 2040
Qy 393 SerAspGlnTrpHis---ValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyPro 411
   ::::| | | | | ::::| | | | | ::::| | | | | ::::| | | | |
Db 2041 CAACATAAAATGAGAATGTTCTTGGTGTTCATTGTTCTCTGGTGGTGGATGTGCCATTCCA 2100
Qy 412 SerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyrAsnVal 430
   ::::| | | | | ::::| | | | | ::::| | | | | ::::| | | | |
Db 2101 AATCGTCCTGGTATTTTGTCCGAGTAGCATATTATGCAAAATGGATACACAAAATT 2157
```

Search completed: December 7, 2005, 00:57:52
Job time : 72 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:49:20 ; Search time 12 Seconds
(without alignments)
208.093 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKPLRPK.....VYTKVSALNWIYNWKAEL 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 41597 seqs, 5740495 residues

Total number of hits satisfying chosen parameters: 41597

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:
1: /cgn2_6/ptodata/2/paa/pct NEW COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06 NEW COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07 NEW COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08 NEW COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US09 NEW COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US10 NEW COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US11 NEW COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US60 NEW COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	461	6	US-10-556-478-3
2	2333	99.6	435	7	US-11-268-554-221
3	2328	99.4	437	7	US-11-268-554-220
4	2297.5	98.1	432	6	US-10-206-921A-330
5	2096	89.5	391	6	US-10-556-478-2
6	1296	55.3	486	6	US-10-556-478-1
7	658	28.1	453	6	US-10-206-921A-64
8	574	24.5	457	8	US-60-732-162-488
9	478	20.4	314	6	US-10-868-184C-4975
10	478	20.4	340	6	US-10-556-478-5
11	475	20.3	423	6	US-10-206-921A-320
12	475	20.3	423	6	US-11-265-762-146
13	467.5	20.0	290	7	US-11-264-096-1294
14	463	19.8	855	7	US-11-268-554-284
15	459.5	19.6	812	7	US-11-108-459-6
16	454.5	19.4	287	7	US-10-556-478-4
17	418.5	17.9	798	7	US-11-108-459-8
18	356	15.2	431	1	PCT-US05-39108-1
19	355.5	15.2	415	7	US-11-166-028-1
20	355.5	15.2	461	7	US-11-172-459-1
21	347	14.8	728	8	US-60-732-162-278
22	343	14.6	261	1	PCT-US05-4045-12
23	323	13.8	699	8	US-60-732-162-276
24	299.5	12.8	723	8	US-60-732-162-1604
25	299.5	12.8	728	6	US-10-555-925-1

ALIGNMENTS

RESULT 1
US-10-556-478-3
; Sequence 3, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diabex, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-556-478-3

Query Match 100.0% Score 2342; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 2.1e-248; Indels 0; Gaps 0;
Matches 435; Conservative 0; Mismatches 0;

QY	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGPIIIALLSLASIIIVVLIKVLIDKYYF	60
DB	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGPIIIALLSLASIIIVVLIKVLIDKYYF	60
QY	61	LCGQPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLSATGN	120
DB	61	LCGQPLHFI PRKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLSATGN	120
QY	121	WFSACFDNPTALATACRQGYSSKPTFRVAVRIGPDQDLDVVEITENSQELRMNSGP	180
DB	121	WFSACFDNPTALATACRQGYSSKPTFRVAVRIGPDQDLDVVEITENSQELRMNSGP	180
QY	181	CLUGSLVSLHCLACGKSLKTRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLIT	240
DB	181	CLUGSLVSLHCLACGKSLKTRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPHWLIT	240

QY 241 AAHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db 241 AAHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
Db 301 SGTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
QY 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435

RESULT 2

US-11-268-554-221
; Sequence 221, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221
; LENGTH: 435
; TYPE: PRP
; ORGANISM: Homo sapiens
US-11-268-554-221

Query Match 99.6%; Score 2333; DB 7; Length 435;
Best Local Similarity 99.8%; Pred. No. 1.9e-247;
Matches 434; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDPSDQPLNSLDVVKPLRKPIMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 60
Db 1 MDPSDQPLNSLDVVKPLRKPIMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 60
QY 61 LCGQPLHPIPRKQLCDGELDCPLGDEBEHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
Db 61 LCGQPLHPIPRKQLCDGELDCPLGDEBEHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
QY 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRNRSSGP 180
Db 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRNRSSGP 180
QY 181 CLSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLTA 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLTA 240
QY 241 AAHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
Db 241 AAHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
Db 301 SGTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
Db 361 QGEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 420
QY 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435

RESULT 3

US-11-268-554-220
; Sequence 220, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 220
; LENGTH: 437
; TYPE: PRP
; ORGANISM: Homo sapiens
US-11-268-554-220

Query Match 99.4%; Score 2328; DB 7; Length 437;
Best Local Similarity 99.8%; Pred. No. 6.6e-247;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVVKPLRKPIMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 61
Db 4 DPDSQPLNSLDVVKPLRKPIMETFRKVGIPPIIIALLSLASIIIVVLIKVLDKYFL 63
QY 62 CGQPLHPIPRKQLCDGELDCPLGDEBEHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 121
Db 64 CGQPLHPIPRKQLCDGELDCPLGDEBEHCVKSPPEGPAVAVRLSKORSTLQVLDSATGN 123
QY 122 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRNRSSGP 181
Db 124 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVDVEITENSQELRNRSSGP 183
QY 182 LSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGKSLKTPRVVVGEEASVDSWPQVSIQYDKQHVCGGSILDPHWLTA 243
QY 242 AHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 301
Db 244 AHCFRKHDTVFNWVVRAGSDKLGSPFLAVAKIIIEFNPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 361
Db 304 GTVRPCLPFFDEELTPATLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADAY 363
QY 362 GEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 421
Db 364 GEVTEKMCAGIPEGGVDTCCGDSGGLMYQSDQHWVGVISWGYCGGPGSTPGVYTKV 423
QY 422 AYLNWIYNWKAEL 435
Db 424 AYLNWIYNWKAEL 437

RESULT 4

US-10-206-921A-330
; Sequence 330, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A

```
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 330
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-921A-330

Query Match      98.1%; Score 2297.5; DB 6; Length 432;
Best Local Similarity 98.8%; Pred. No. 1.4e-243;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY  2 DPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPPIIIALLSLASIIIVVILKVIDKYFL 61
DB  4 DPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPPIIIALLSLASIIIVVILKVIDKYFL 63

QY  62 CGQPLHFIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 121
DB  64 CGQPLHFIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGNW 123

QY  122 FSACFDNFTEALAEACRQMGYSKPTFRAVEIGPQDQDLDVVEITENSQELMRNSSGPC 181
DB  124 FSACFDNFTEALAEACRQMGYS-----RAVEIGPQDQDLDVVEITENSQELMRNSSGPC 178

QY  182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHVLTA 241
DB  179 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHVLTA 238

QY  242 AHCPRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLQPLTFS 301
DB  239 AHCPRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLQPLTFS 298

QY  302 GTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQ 361
DB  299 GTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAYQ 358

QY  362 GVETKMKCAGIPEGGVDTCCQDSGGPLMYQSDQHWVGVISWGYGCGSPTPGVYTKVS 421
DB  359 GVETKMKCAGIPEGGVDTCCQDSGGPLMYQSDQHWVGVISWGYGCGSPTPGVYTKVS 418

QY  422 AYLNWIYNWKAE 435
DB  419 AYLNWIYNWKAE 432

RESULT 5
US-10-556-478-2
; Sequence 2, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; PRIOR FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 391

; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-556-478-2

Query Match      89.5%; Score 2096; DB 6; Length 391;
Best Local Similarity 99.7%; Pred. No. 1.4e-221;
Matches 384; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  51 IKVILDKYFLCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRST 110
DB  1 MKVILDKYFLCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRST 60

QY  111 LQVLD SATGNWFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPQDQDLDVVEITENSQ 170
DB  61 LQVLD SATGNWFSACFDNFTEALAEACRQMGYSKPTFRAVEIGPQDQDLDVVEITENSQ 120

QY  171 ELRNRSSGPGCLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGG 230
DB  121 ELRNRSSGPGCLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGG 180

QY  231 SILDPHWLVTAACHFRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIA 290
DB  181 SILDPHWLVTAACHFRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIA 240

QY  291 LMKLQPLTFSGTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVID 350
DB  241 LMKLQPLTFSGTVRPICLPFFDEELTPATPLIIGWFTKQNGKMSDILLQASVQVID 300

QY  351 STRCNADDAYQGVETKMKCAGIPEGGVDTCCQDSGGPLMYQSDQHWVGVISWGYGCGG 410
DB  301 STRCNADDAYQGVETKMKCAGIPEGGVDTCCQDSGGPLMYQSDQHWVGVISWGYGCGG 360

QY  411 PSTGVTYTKVSAYLNWIYNWKAE 435
DB  361 PSTGVTYTKVSAYLNWIYNWKAE 385

RESULT 6
US-10-556-478-1
; Sequence 1, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-556-478-1

Query Match      55.3%; Score 1296; DB 6; Length 486;
Best Local Similarity 60.1%; Pred. No. 7.9e-134;
Matches 270; Conservative 21; Mismatches 56; Indels 102; Gaps 9;

QY  48 VVILKVIDKYFLCGQPLHFIPRKQCDGELDCPLGEDEEHCVKSPFEGPAVAVR---- 103
DB  69 MAIRYIADKHNWLG-----CP-----KERAELSMLEGAVIDIRYGV 107
```

QY 104 ---LSKDRSTLOV-----LDS 116
Db 108 RIAYSKDFETLKVDLSKLPMLKMFEDRLCHTKYTLNGDHVTHPDRLYDALDVVLYMDP 167
QY 117 ATGNWFS--ACFNFTAL--AETACROMGYSKP-----TPRAVEIGDQDOLDVVEIT 166
Db 168 MCLDAPFKLVCFKKRIEAIPOIDIKLKSSKIYIAIPLQGWQATPGGDDHPPKSDL-----VP 223
QY 167 ENSQELMRNNSGPGCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQVSIQYDKOH 226
Db 224 RHNQ-----TSLYKAGFENLYFQGVVGGEASVDSWPQVSIQYDKOH 267
QY 227 VCGGSLDHPHWLTAHCFRKHDTDFNWKVRAGSDKLGSPFSLAVAKIIIBFNPMYKPD 286
Db 268 VCGGSLDHPHWLTAHCFRKHDTDFNWKVRAGSDKLGSPFSLAVAKIIIBFNPMYKPD 327
QY 287 NDIALMKLOPPLTFSGTVRPICLPFFDELTATPLWIIIGWFTKQNGKMSDILLQASV 346
Db 328 NDIALMKLOPPLTFSGTVRPICLPFFDELTATPLWIIIGWFTKQNGKMSDILLQASV 387
QY 347 QVIDSTRCNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQWHVVGIVSWGY 406
Db 388 QVIDSTRCNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQWHVVGIVSWGY 447
QY 407 GCGGPGTGPVYTKVSAYLWNIYVNWKAEL 435
Db 448 GCGGPGTGPVYTKVSAYLWNIYVNWKAEL 476

RESULT 7
US-10-206-921A-64
; Sequence 64, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: 39780-3430R1C515
; CURRENT APPLICATION NUMBER: US/10/206,921A
; PRIOR FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: US 10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/232,887
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 64
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-206-921A-64

Query Match 28.1%; Score 658; DB 6; Length 453;
Best Local Similarity 35.7%; Pred. No. 5.1e-64;
Matches 158; Conservative 75; Mismatches 146; Indels 64; Gaps 18;

QY 22 IPMTFRKVGPIPIIALLSLASIIIVVLVKILD---KY-----YFLCGQPLHFIPRKQL 74
Db 44 LPLRFFPIIIGIITALILA-----IGLIHFDCSGKRCRSSFKC---IELIAR---91
QY 75 CDGBLDCPLGEDEHCVKSPFEGPAVAVLSKDRSTLOVLDATGNWFSACFNFTAL 134
Db 92 CDGVSDCKGDEDEYRC-----VRVGQNAVLOVFTAA--SNKTCSDDDMKGHYA 138

QY 135 ETACROMGYSS-----KPTTFAVEIGDQDOLDVVBVITENSQELMRNNSGPGCL 182
Db 139 NVACAQLGFPYSSVSDNLRVSSLEGQFREBFVSIDHLLPDDKVTAALHHSVVVREG---CA 195
QY 183 SGLSVLSHCLACG-KSLKTPRVVGGEASVDSWPQVSIQYDKOHVCGGSLDHPHWLTA 241
Db 196 SGHVVTIQCACGHRRGYSRIRVCGNMSLISQWPQASLOFQGYHLGCGSVITPLWITA 255
QY 242 AHCFRKHDTDFN--NWKVRAGSDKL--GSPFSLAVAKIIIBFNPMY-PK--DNDIALMK 293
Db 256 AHCV---YDLYLPKSWTIQVLSLNDNPAPSHLVEKIV---YHSKYKPKRLGNDIALMK 309
QY 294 LQPLTFSGTVRPICLPFFDELTATPLWIIIGWFTKQNGKMSDILLQASVQVIDSTR 353
Db 310 LAGPLTFNEMIQVCLPNSSENFDPGVKVCWTSGWGAT-EDGGDASPVLNHAAVPLISNKI 368
QY 354 CNADDAAYQGEVTEKMCAGIPGGVDTCQDGGGGLMYQSDQ-WHVVGIVSWGVCGGPS 412
Db 369 CNHRDVTGGIISPSMLCAGYLTGGVDSQDGGGLVQCERRLWKLVGATSGFGTCAEVN 428
QY 413 TPGVYTKVSAYLWNIYVNWKAEL 435
Db 429 KPGVYTRVTSFLDNIHQMERDL 451

RESULT 8
US-60-732-162-488
; Sequence 488, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Berdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60/732,162
; CURRENT FILING DATE: 2005-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 488
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Homosapiens
US-60-732-162-488

Query Match 24.5%; Score 574; DB 8; Length 457;
Best Local Similarity 32.2%; Pred. No. 8.2e-55;
Matches 129; Conservative 70; Mismatches 151; Indels 50; Gaps 12;

QY 60 FLC---GQPLHFIPRKQLCDGE--LDCPLGEDEHCVKSPFEGPAVAVRLSKDRSTLOVL 114
Db 71 YLCPAASQPI-----SGTLQDEEITLSCSEASAEALLPALPK--TVSFRINSEDFLLBAQ 124
QY 115 DSATGNWFSACFNFTALAEATACRQMGYSKPTTFAVEIGDQDOLDVVBVITENSQE--- 171
Db 125 VRDQPRLLVLCHEGWSPALGLQICWSLGH-----LRLTHKGVNITDKLNSQSFBA 176
QY 172 -----LMRNNSGPGCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQV 218
Db 177 QLSPLRGLFBEAWQPRNN---CTSGQVVSURCSECGARPLASRIVGGQSVAPGRWPQA 233
QY 219 STQYDKOHVCGGSLDHPHWLTAHCFR--KHTDVFNWVKVPAAGSKLGSFPFSLAVAKIII 276

Db 234 SVALGFRHTECGSVLAPRWVTAACHMHSFRLARLSSWRVHAG---LVSHSAVRPHQAL 290
Qy 277 IEF---NPMYKDN---DIALMKLQPLTFSTGTVRPICLPFFDEBELTATPLMIIGWGFT 330
Db 291 VERIIPHPYSAQNHDYDVALRLQTLNFSDTVGAVCLPAKEQHFPKGSRCWVGWGT 350
Qy 331 KONGKMSDILLQASVQVIDSTRCNADAYQGEVTEKMKCAGIPGGGVDTCCGDSGGPLM 390
Db 351 HPSHTYSDDLQDTVPLFSTQLCNSSCVSYGALTFRMLCAGYLDGRADACQDGGGGLV 410
Qy 391 Y-OSDQWVGVSVKSGCGGPGSTPGVYTKVSAYLNLWYN 429
Db 411 CPDGTWRLGVVSWGRACAEHPHGVYAKVAEFLDWIHD 450

RESULT 9

US-10-868-184C-4975
; Sequence 4975, Application US/10868184C
; GENERAL INFORMATION:
; APPLICANT: Rosen, et. al
; TITLE OF INVENTION: Human Secreted Proteins
; FILE REFERENCE: PS805
; CURRENT APPLICATION NUMBER: US/10/868,184C
; PRIOR FILING DATE: 2004-06-16
; PRIOR APPLICATION NUMBER: 60/278,650
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/833,245
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US01/11988
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: PCT/US00/06043
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06012
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06058
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06044
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06059
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06042
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: PCT/US00/06014
; PRIOR FILING DATE: 2000-03-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 13046
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4975
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-868-184C-4975

Query Match 20.4%; Score 478; DB 6; Length 314;
Best Local Similarity 37.8%; Pred. No. 1.5e-44;
Matches 101; Conservative 48; Mismatches 86; Indels 32; Gaps 8;

Qy 178 SGPCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPW 237
Db 30 SGP-----CGRRVITSRIVGGEDELGRWPMQGSRLRLWDSHVCGVSLSHRW 76
Qy 238 VLTAACHCRKHTDVFN---WKVRAGSDKLGSPFSLAVAK-----IIIEFNPMYPKDN 287
Db 77 ALTAACHCFETYDLSLDPGSMWVQFG--QLTSMPSFWSLQAYYTRYFVSNVILSPRYLGS 134
Qy 288 --DIALMKLQPLTFSTGTVRPICLPFFDEBELTATPLMIIGWFTKQNGKMSDILLQ-A 344
Db 135 PYDIALVKLSAPVTVTKHIQICLQASTFEFENRTDCWVTGWYKIKEDALPSPHTLQEV 194
Qy 345 SVQVIDSTRCN---ADDAVQGEVTEKMKCAGIPGGGVDTCCGDSGGPL-MYQSDQWVVVG 400
Db 195 QVAIIINSMCNHLFLKYSFRKDIIFGDMVCAGNAQGGKDACFGDSGGGLACNKNGLWYQIG 254

Qy 401 IVSNVGYCGGPGSTPGVYTKVSAYLNLWI 427
Db 255 VVSWGVCGRNRPVGTNTNISHHFEWI 281

RESULT 10

US-10-556-478-5
; Sequence 5, Application US/10556478
; GENERAL INFORMATION:
; APPLICANT: diadexus, Inc.
; APPLICANT: Pilkington, Glenn
; APPLICANT: Keller, Gilbert-Andre
; APPLICANT: Li, Wenlu
; APPLICANT: Corral, Laura
; APPLICANT: Simon, Iris
; TITLE OF INVENTION: Ovr115 Antibody Compositions and Methods of Use
; FILE REFERENCE: DEX-0484
; CURRENT APPLICATION NUMBER: US/10/556,478
; CURRENT FILING DATE: 2005-11-14
; PRIOR APPLICATION NUMBER: US 60/559,730
; PRIOR FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: US 60/471,068
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-556-478-5

Query Match 20.4%; Score 478; DB 6; Length 340;
Best Local Similarity 37.8%; Pred. No. 1.7e-44;
Matches 101; Conservative 48; Mismatches 86; Indels 32; Gaps 8;

Qy 178 SGPCLSGSLVSLHCLACGKSLKTPRVVGGEASVDSWPQVSIQYDKQHVCGGSLDHPW 237
Db 30 SGP-----CGRRVITSRIVGGEDELGRWPMQGSRLRLWDSHVCGVSLSHRW 76
Qy 238 VLTAACHCRKHTDVFN---WKVRAGSDKLGSPFSLAVAK-----IIIEFNPMYPKDN 287
Db 77 ALTAACHCFETYDLSLDPGSMWVQFG--QLTSMPSFWSLQAYYTRYFVSNVILSPRYLGS 134
Qy 288 --DIALMKLQPLTFSTGTVRPICLPFFDEBELTATPLMIIGWFTKQNGKMSDILLQ-A 344
Db 135 PYDIALVKLSAPVTVTKHIQICLQASTFEFENRTDCWVTGWYKIKEDALPSPHTLQEV 194
Qy 345 SVQVIDSTRCN---ADDAVQGEVTEKMKCAGIPGGGVDTCCGDSGGPL-MYQSDQWVVVG 400
Db 195 QVAIIINSMCNHLFLKYSFRKDIIFGDMVCAGNAQGGKDACFGDSGGGLACNKNGLWYQIG 254
Qy 401 IVSNVGYCGGPGSTPGVYTKVSAYLNLWI 427
Db 255 VVSWGVCGRNRPVGTNTNISHHFEWI 281

RESULT 11

US-10-206-921A-320
; Sequence 320, Application US/10206921A
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

```
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: 39780-3430R1C515
/ CURRENT APPLICATION NUMBER: US/10/206,921A
/ CURRENT FILING DATE: 2002-07-26
/ PRIOR APPLICATION NUMBER: US 10/052,586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US 60/232,887
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 320
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ ORGANISM: Homo Sapien
US-10-206-921A-320

Query Match      20.3%; Score 475; DB 6; Length 423;
Best Local Similarity 39.6%; Pred. No. 5e-44;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLACG----KSL-KTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWWLTAHC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 HC--CGTRRSKTLGQSLRIVGGTEVEEGEWPQASLQWDGSHRCGATLINATLWLSAAHC 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 245 FRKHTDVFNKVRAGSDKLGSPFLAVAKIIIEFNPMYFKDNDIALMKLQFLPFTSGTV 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 FTYKNPARTWASFGVTIKPSKMKGLRRIIVHEKYKHPSHDYDISLAELSSPVYTNV 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 305 RPICLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVVIDSTRCNADDAYQGEV 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 HRVCLPDASYEQFGDVMEVTFGALK-NDGYSQNHRLRQAQVTLIDATTCNEPQYNDAI 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 365 TEKMCAGIPEGGVDTCCGDSGGPLMYQS--DQHVVGIVSWGCGGPGSTPGVYTKVSA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TPRMLCAGSLEGTACQGDGSGPLVSSDARDIWLVLGIVSWGDECAKPNKPGVYTRVTA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 423 YLNWI 427
    |||
Db 413 LRDWI 417
    |||

RESULT 12
US-11-265-762-146
/ Sequence 146, Application US/11265762
/ GENERAL INFORMATION:
/ APPLICANT: McCarthy, Sean A.
/ TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
/ FILE REFERENCE: 7853-227-999
/ CURRENT APPLICATION NUMBER: US/11/265,762
/ CURRENT FILING DATE: 2005-11-02
/ PRIOR APPLICATION NUMBER: US/09/796,753
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 09/183,175
/ PRIOR FILING DATE: 1998-10-30
/ PRIOR APPLICATION NUMBER: 09/223,094
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/223,546
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/224,246
/ PRIOR FILING DATE: 1998-12-30
/ PRIOR APPLICATION NUMBER: 09/259,388
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/122,458
/ PRIOR FILING DATE: 1999-03-01
/ PRIOR APPLICATION NUMBER: 09/312,359
/ PRIOR FILING DATE: 1999-05-14
/ PRIOR APPLICATION NUMBER: 09/336,536
/ PRIOR FILING DATE: 1999-06-18
/ PRIOR APPLICATION NUMBER: 09/342,687
/ PRIOR FILING DATE: 1999-06-29
/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: 39780-3430R1C515
/ CURRENT APPLICATION NUMBER: US/10/206,921A
/ CURRENT FILING DATE: 2002-07-26
/ PRIOR APPLICATION NUMBER: US 10/052,586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: PCT/US01/06520
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US 60/232,887
/ PRIOR FILING DATE: 2000-09-15
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 320
/ LENGTH: 423
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
/ ORGANISM: Homo Sapien
US-10-206-921A-320

Query Match      20.3%; Score 475; DB 6; Length 423;
Best Local Similarity 39.6%; Pred. No. 5e-44;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLACG----KSL-KTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWWLTAHC 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 HC--CGTRRSKTLGQSLRIVGGTEVEEGEWPQASLQWDGSHRCGATLINATLWLSAAHC 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 245 FRKHTDVFNKVRAGSDKLGSPFLAVAKIIIEFNPMYFKDNDIALMKLQFLPFTSGTV 304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 FTYKNPARTWASFGVTIKPSKMKGLRRIIVHEKYKHPSHDYDISLAELSSPVYTNV 293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 305 RPICLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVVIDSTRCNADDAYQGEV 364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 294 HRVCLPDASYEQFGDVMEVTFGALK-NDGYSQNHRLRQAQVTLIDATTCNEPQYNDAI 352
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 365 TEKMCAGIPEGGVDTCCGDSGGPLMYQS--DQHVVGIVSWGCGGPGSTPGVYTKVSA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TPRMLCAGSLEGTACQGDGSGPLVSSDARDIWLVLGIVSWGDECAKPNKPGVYTRVTA 412
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 423 YLNWI 427
    |||
Db 413 LRDWI 417
    |||

RESULT 13
US-11-264-096-1294
/ Sequence 1294, Application US/11264096
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PF546D1
/ CURRENT APPLICATION NUMBER: US/11/264,096
/ CURRENT FILING DATE: 2005-11-02
/ PRIOR APPLICATION NUMBER: 09/833,245
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229,358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256,931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199,384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 2267
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1294
/ LENGTH: 290
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ ORGANISM: Homo sapiens
US-11-264-096-1294

Query Match      20.0%; Score 467.5; DB 7; Length 290;
Best Local Similarity 37.1%; Pred. No. 1.8e-43;
Matches 99; Conservative 49; Mismatches 92; Indels 27; Gaps 9;

QY 184 GSLVSLHCLACGSKLKTPTRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWWLTAH 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16 GSGRAKAATAACGPRMLNRMVGGQDTQEGEWPQVSIQRNGSHFCGGSLLIAEQWLTAH 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 244 CERKHTDVFNKVRAGSDKL-GSPFLAVAKIIIEFNPMY---PKDNDIALMKLQFLPT 299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 CFRNTSETSLYQVLLGARQLVQPGPHAMYARVQVESNPLYQGTASSADVALVEAPVP 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 300 FSGTVRPICLPFFDEELTPATPL--WIIGWFTKQNGKMSD-----ILLQASVQVIDS 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 136 FTNYILPCLP--DPSVIFETGMNCWVTGWSGSPSE-----EDLLPEPILQKLAIPIDT 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Qy 352 TRCN-----ADDAVQGE-VTEKMMKAGIPBGGVDTCCGDSGGPIMYQSDQ-WHVVGIVS 403
Db 189 PKCNLLYSKDTFQGYQPKTIKNDMLCAGFEFGKDKACKGGLVCLVGLGQSLQAGVIS 248
Qy 404 WGYCGGSPSTPGVYTKVSAYLNWIYV 430
Db 249 WGECCARQNRGVYIRVTAHHNWIHRI 275

RESULT 14

US-11-268-554-284
; Sequence 284, Application US/11268554
; GENERAL INFORMATION:
; APPLICANT: DOMON, Bruno et al.
; TITLE OF INVENTION: PANCREATIC CANCER TARGETS AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CL001557
; CURRENT APPLICATION NUMBER: US/11/268,554
; CURRENT FILING DATE: 2005-11-08
; NUMBER OF SEQ ID NOS: 1004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-268-554-284

Query Match 19.8%; Score 463; DB 7; Length 855;
Best Local Similarity 31.5%; Pred. No. 3.1e-42;
Matches 123; Conservative 63; Mismatches 151; Indels 54; Gaps 18;

Qy 58 YFLCG-----QPLHPIPRKQICDGLDCPLGDEHCVKSPFEGPAAVAVRUSKORSTLOV 113
Db 492 HQFTCKNFKCKPLFW-----CDSVNDGDNSEQGC-----SCPAQTFRCGNGKC---L 538
Qy 114 LDSATGNWFSACFNFTALAEACROMGY--SSKPTFAVEIGPDQDLDVVEITENSQE 171
Db 539 SKSQOCNGKDCGDSDEA---SCPKNVVTCTKHTYRC-----LNGCLSKGNPE 586
Qy 172 LRMNSSGFCLSGLSVSLHCLACG--KSLKTPRVVVGEEASVDSWPQVSIQ-YDKQHV 228
Db 587 C---DGKEDCSGDS-DEKDC-DCGLRSTRQARVVGVDADGEWFWQVSLHALGQGHIC 641
Qy 229 GGSILDPHVLTAACHF-----RKHTDVFNKVRAG--SDKLGSPPSLAVAKIIIEFNP 281
Db 642 GASLISPMLVSAACHYIDDRFRYSPTQWTFGLHDQSQRSAPGVQERRKRLIISHP 701
Qy 282 MPK---DNDIALMKLQPLTFSGTVRPICLPFFDEELTPATPLWIIWGFTKQNGGKMS 338
Db 702 FFNDFTFDYDIALLELEKPAEYSSMVRPICLPDASHVFPAGKAIWVTGWGHT-QYGGTGA 760
Qy 339 DILLQASVQVIDTRCNADDAVQGEVTEKMMKAGIPBGGVDTCCGDSGGPIMYQSDQWHV 398
Db 761 LILQGEIRVINQTC--ENLLPQOITPRMCMVGLSGVDSCQDGGPLSSVEADGRI 818
Qy 399 --VGIVSMGYCGGSPSTPGVYTKVSAYLNWI 427
Db 819 FQAGVWSGDCGAQRNKGVVYTRLPFLFRDWI 849

RESULT 15

US-11-108-459-6
; Sequence 6, Application US/11108459
; GENERAL INFORMATION:
; APPLICANT: Zaa, Aimee
; APPLICANT: Schwartz, David A.
; APPLICANT: Peltz, Gary
; TITLE OF INVENTION: Polymorphic Plasminogen Genes and Uses Thereof
; FILE REFERENCE: DUKE-09785
; CURRENT APPLICATION NUMBER: US/11/108,459
; CURRENT FILING DATE: 2005-04-18
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in version 3.3

; SEQ ID NO 6
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-108-459-6

Query Match 19.6%; Score 459.5; DB 7; Length 812;
Best Local Similarity 39.8%; Pred. No. 7e-42;
Matches 103; Conservative 43; Mismatches 88; Indels 25; Gaps 8;
Qy 181 CLSGSLVSLHCLACGKSLKTP-----RVVGEEASVDSWPQVSI--QYDQHVCGGSIL 233
Db 560 CASAS-----SPEGCKPQVEPKCPGRVVGCVANPHSWPQISLRTFTTQGHFCGGTLI 614
Qy 234 DPHWVLTAAHCFRKHTDVFNMKVRAGSDK---LGS-FPSLAVAKIIIEFNPMYPKNDI 289
Db 615 APEWVLTAAHCLEKSSRPEFYKVLGAHEEYIRGSDVQEISVAKLI-----LEPNNRDI 668
Qy 290 ALMKLQPLTFSGTVRPICLPFFDEELTPATPLWIIWGFTKQNGKMSDILLQASVOVI 349
Db 669 ALLKLSRPATITDKVIPACLPSPNYVADRITCITGWGETQGTFG--AGRLKSAQLPVI 726
Qy 350 DSTRCNADDAVQGEVTEKMMKAGIPBGGVDTCCGDSGGPLM-YOSDQWHVVGIVSMGYGC 408
Db 727 ENKVCNRVEYLNRRVKSTELCAGQLAGVDSQGDSCGGLVCFPKDKYILQGVTSWGLGC 786
Qy 409 GGPSTPGVYTKVSAYLNWI 427
Db 787 ARPKNKPGVYVRSRFDWI 805

Search completed: December 5, 2005, 14:04:42
Job time : 14 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:42:04 ; Search time 231 Seconds
(without alignments)
1328.592 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVPLRKP.....VTKVSAYLNIYNNWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705529306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2337	99.8	437	1	TMPS4_HUMAN
2	1844	78.7	435	1	TMPS4_MOUSE
3	762	32.5	388	2	Q4RR7_TETNG
4	705.5	30.1	543	1	TMPSD_MOUSE
5	699.5	29.9	359	2	Q4SPG0_TETNG
6	692.5	29.6	581	1	TMPSD_HUMAN
7	691.5	29.5	486	2	Q5PRAG_BRARE
8	688	29.4	453	2	Q812A6_MOUSE
9	687	29.3	453	1	TMPS3_MOUSE
10	686.5	29.3	767	2	Q9DGR2_XENLA
11	680.5	29.1	722	2	Q6NUF5_XENLA
12	676.5	28.9	492	2	Q6GTK7_HUMAN
13	673.5	28.8	492	1	TMPS2_HUMAN
14	670.5	28.6	492	2	Q9FT73_HUMAN
15	660.5	28.2	454	1	TMPS3_HUMAN
16	660.5	28.2	538	2	Q5USC7_HUMAN
17	658	28.1	490	2	Q6P7D7_RAT
18	648.5	27.7	490	1	TMPS2_MOUSE
19	648	27.7	490	2	Q7TN04_MOUSE
20	646	27.6	490	2	Q920K3_RAT
21	636.5	27.2	445	2	Q8CJ17_RAT
22	634	27.1	439	2	Q5RDX7_PONPY
23	631.5	27.0	455	2	Q8CDR0_MOUSE
24	628.5	26.8	455	2	TMPS5_MOUSE
25	612	26.1	371	2	Q8CJ16_RAT
26	585.5	25.0	417	2	Q5R5E8_PONPY
27	584.5	25.0	417	1	HEPS_HUMAN
28	582.5	24.9	1019	1	ENTK_BOVIN
29	578	24.7	436	1	HEPS_MOUSE
30	574	24.5	457	1	TMPS5_HUMAN
31	571	24.4	1035	1	ENTK_BOVIN

32	568.5	24.3	416	1	HEPS_RAT
33	568	24.3	418	1	TM1D_HUMAN
34	552.5	23.6	1069	1	ENTK_MOUSE
35	552	23.6	1034	1	ENTK_PIG
36	550.5	23.5	799	1	TMPS6_MOUSE
37	550.5	23.5	799	2	Q6PF94_MOUSE
38	545	23.3	777	2	Q8CAN9_MOUSE
39	536	22.9	417	1	TM1D_RAT
40	531	22.7	802	1	TMPS6_HUMAN
41	523	22.3	417	1	TM1D_MOUSE
42	518.5	22.1	418	2	Q6IE15_RAT
43	514.5	22.0	730	2	Q4RHT0_TETNG
44	511	21.8	638	1	KLKBI_RAT
45	511	21.8	638	2	Q5FVS2_RAT

RESULT 1

ID TMPS4_HUMAN STANDARD; PRT; 437 AA.

AC Q9NRS4; OSXK06; Q6UX37; Q9NZAS;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Transmembrane protease, serine 4 (EC 3.4.21.-) (Membrane-type serine protease 2) (MT-SP2).

GN Name=TMPSR34; Synonyms=TMPSR33; ORFNames=UNQ776/PRO1570;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.

OC NCBI_TaxID=9606; ..

RN [1]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).

RC TISSUE=Pancreatic carcinoma;

RX MEDLINE=20283276; PubMed=10825129;

RA Wallrapp C., Haehnel S., Mueller-Pillasch F., Burghardt B., Iwamura T., Ruthenburger M., Lerch M.M., Adler G., Gress T.M.;

RT "A novel transmembrane serine protease (TMPSR33) overexpressed in pancreatic cancer.";

RT Cancer Res. 60:2602-2606 (2000).

RL [2]

RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).

RA Smeekens S.S., Lorimer D.D., Wang E., Hou J., Linnevers C.;

RT "MT-SP2, a novel type II membrane serine protease expressed in trachea, colon, and small intestine: identification, cloning, and chromosomal localization.";

RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).

TISSUE=Colon;

RC PubMed=14702039; DOI=10.1038/ng1285;

| RA | Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiyama S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshioka S., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Moniyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu P., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., | |

RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
RA Kawabata N., Hikiiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani A., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs. Genet. 36:40-45 (2004).
RL Nat. Genet. 36:40-45 (2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
RA Chen J., Chow B., Chul C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith Y., Stinson J., Vagts A.,
RA Vandlen R.L., Watanabe C., Wiedand D., Woods K., Xie M.-H.,
RA Yaneura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
RA Wood W.I., Godowski P.J., Gray A.M.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2265-2270(2003).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 3).
RC TISSUE=Ovary, and Pancreas;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues Y., Bouffard G.G.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Smolius D.E.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
-!- FUNCTION: Probable protease. Seems to be capable of activating
CC ENAC (By similarity).
-!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q9NRS4-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9NRS4-2; Sequence=VSP_013117;
CC Note=No experimental confirmation available;
CC Name=3;
CC IsoId=Q9NRS4-3; Sequence=VSP_013116;
CC Note=No experimental confirmation available;
CC TISSUE SPECIFICITY: High levels in pancreatic, gastric, colorectal
CC and ampullary cancer. Very weak expression in normal
CC gastrointestinal and urogenital tract.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a

frameshift in position 31.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

CC EMBL; AF179224; AAF74526.1; -; mRNA.
CC EMBL; AF216312; AAF31436.1; ALT FRAME; mRNA.
CC EMBL; AK172766; BADI8749.1; -; mRNA.
CC EMBL; AY358530; AAQ88894.1; -; mRNA.
CC EMBL; BC004855; AAH04855.1; -; mRNA.
CC EMBL; BC011703; AAH11703.1; -; mRNA.
CC HSSP; P00760; IEZX.
CC MEROPS; S01.034; -.
DR EMBL; ENSG00000137648; Homo sapiens.
DR HGNC; HGNC:11878; TMRSS4.
DR H-InvDB; HIX0010173; -.
DR MIM; 606565; -.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR InterPro; IPR001272; LDL receptor A.
DR InterPro; IPR001254; Peptidase S1-S6.
DR InterPro; IPR001314; Peptidase S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00200; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLRA_1; FALSE NEG.
DR PROSITE; PS00568; LDLRA_2; FALSE NEG.
DR PROSITE; PS00420; SRCR_1; FALSE NEG.
DR PROSITE; PS0287; SRCR_2; 1.
DR PROSITE; PS0240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Alternative splicing; Glycoprotein; Hydrolase; Protease;
KW Serine protease; Signal-anchor; Transmembrane.
FT TOPO_DOM 1 32 Cytoplasmic (Potential).
FT TRANSMEM 33 53 Signal-anchor for type II membrane
FT TOPO_DOM 54 437 Extracellular (Potential).
FT DOMAIN 61 93 LDL-receptor class A.
FT DOMAIN 94 204 SRCR.
FT DOMAIN 205 434 Peptidase S1.
FT ACT_SITE 245 245 Charge relay system (By similarity).
FT ACT_SITE 290 290 Charge relay system (By similarity).
FT ACT_SITE 387 387 Charge relay system (By similarity).
FT SITE 204 205 Cleavage (Potential).
FT CARBOHYD 130 130 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
FT DISULFID 64 83 By similarity.
FT DISULFID 77 92 By similarity.
FT DISULFID 127 183 By similarity.
FT DISULFID 140 193 By similarity.
FT DISULFID 196 310 By similarity.
FT DISULFID 230 246 By similarity.
FT DISULFID 356 372 By similarity.
FT DISULFID 383 410 By similarity.
FT VARSPPLIC 1 3 MLQ -> M (in isoform 3).
FT VARSPPLIC 147 151 /FTID=VSP_013116.
FT CONFLICT 2 2 Missing (in isoform 2).
FT CONFLICT 2 2 /FTID=VSP_013117;
SQ SEQUENCE 437 AA; 48204 MW; 351B2FD4A8657B12 CRC64;
Query Match 99.8%; Score 2337; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 3.3e-190;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 61
DB 4 DPDSQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 63
QY 62 CGQPLHFTPRKQCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLDLSATGNW 121
DB 64 CGQPLHFTPRKQCDGELDCPLGEDEEHCVKSPFEGPAPAVRLSKDRSTLQVLDLSATGNW 123
QY 122 FSACFDNTEALAEATACRGMYSKPTFRVVEIGPDQDLVDVEITENSQELMRNNSGPC 181
DB 124 FSACFDNTEALAEATACRGMYSKPTFRVVEIGPDQDLVDVEITENSQELMRNNSGPC 183
QY 182 LSGSLVSLHCLACGSLKTPRVGGEASVDSWPMQVSIQVDKHVCGSITLDHWHVLT 241
DB 184 LSGSLVSLHCLACGSLKTPRVGGEASVDSWPMQVSIQVDKHVCGSITLDHWHVLT 243
QY 242 AHCFRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTFS 301
DB 244 AHCFRKHTDVFNKVRAGSKLGSPPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTFS 303
QY 302 GTVRPILCPFFDEELTPATPLMIIGWFTKQNGKRWSDILLQASVQVIDSTRCNADDAYQ 361
DB 304 GTVRPILCPFFDEELTPATPLMIIGWFTKQNGKRWSDILLQASVQVIDSTRCNADDAYQ 363
QY 362 GEVTEKMKCAGIPGEGVDTCCGDSGGLMYQSDQHHVVGIVSWGCGGSPTPGYTVKVS 421
DB 364 GEVTEKMKCAGIPGEGVDTCCGDSGGLMYQSDQHHVVGIVSWGCGGSPTPGYTVKVS 423
QY 422 AYLNWYNNWKAEL 435
DB 424 AYLNWYNNWKAEL 437

RESULT 2
ID TMS4 MOUSE
AC Q8VCA5;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transmembrane protease, serine 4 (EC 3.4.21.-) (Channel-activating
protease 2) (mCAP2).
DE Name=tmprss4; Synonyms=Cap2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22144321; PubMed=12149280;
RA Vuagniaux G., Vallet V., Jaeger N.F., Hummler E., Rossier B.C.;
RT "Synergistic activation of ENaC by three membrane-bound channel-
activating serine proteases (mCAP1, mCAP2, and mCAP3) and serum- and
glucocorticoid-regulated kinase (Sgk1) in Xenopus oocytes.";
RL J. Gen. Physiol. 120:191-201(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probable protease. Seems to be capable of activating
ENaC.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
CC EMBL: AY043240; AAK85307.1; -; mRNA.
CC EMBL: BC021368; AAH21368.1; -; mRNA.
CC HSSP: P00760; 1EZX.
CC MEROPS: S01.034; -;
CC Ensembl: ENSMUSG00000032091; Mus musculus.
CC MGI: MGI:2384877; Tmprss4.
CC InterPro: IPR002172; LDL_receptor_A.
CC InterPro: IPR001254; Peptidase_S1_S6.
CC InterPro: IPR001314; Peptidase_S1A.
CC InterPro: IPR001190; S1cr_receptor.
CC Pfam: PF00089; Trypsin; 1.
CC PRINTS: PR00722; CHYMOTRYPSIN.
CC SMART: SM00192; LDLA; 1.
CC SMART: SM00202; SR; 1.
CC SMART: SM00202; Tryp_Src; 1.
CC PROSITE: PS01209; LDLRA_1; FALSE_NEG.
CC PROSITE: PS00068; LDLRA_2; FALSE_NEG.
CC PROSITE: PS00420; SRCR_1; FALSE_NEG.
CC PROSITE: PS02087; SRCR_2; 1.
CC PROSITE: PS50240; TRYPSIN_DOM; 1.
CC PROSITE: PS50240; TRYPSIN_HIS; 1.
CC PROSITE: PS00134; TRYPSIN_SER; 1.
CC PROSITE: PS00135; TRYPSIN_SER; 1.
KW Glycoprotein; Hydrolase; Protease; Serine protease; Signal-anchor;
KW Transmembrane.
KW TOPO_DOM 1 30 Cytoplasmic (Potential).
KW TRANSMEM 31 51 Signal-anchor for type II membrane
FT TOPO_DOM 1 30 Cytoplasmic (Potential).
FT TRANSMEM 31 51 Protein (Potential).
FT TOPO_DOM 52 435 Extracellular (Potential).
FT DOMAIN 59 101 LDL-receptor class A.
FT DOMAIN 102 202 SRCR.
FT DOMAIN 203 432 Peptidase S1.
FT ACT_SITE 243 243 Charge relay system (By similarity).
FT ACT_SITE 288 288 Charge relay system (By similarity).
FT ACT_SITE 385 385 Charge relay system (By similarity).
FT SITE 202 203 Cleavage (Potential).
FT CARBOHYD 128 128 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 176 176 N-linked (GlcNAc...) (Potential).
FT DISULFID 62 81 By similarity.
FT DISULFID 75 90 By similarity.
FT DISULFID 125 181 By similarity.
FT DISULFID 138 191 By similarity.
FT DISULFID 194 308 By similarity.
FT DISULFID 228 244 By similarity.
FT DISULFID 354 370 By similarity.
FT DISULFID 381 408 By similarity.
SQ SEQUENCE 435 AA; 47496 MW; DC52E45A43E01369 CRC64;
Query Match 78.7%; Score 1844; DB 1; Length 435;
Best Local Similarity 77.2%; Pred. No. 3.2e-148;
Matches 336; Conservative 43; Mismatches 56; Indels 0; Gaps 0;
QY 1 MDPSDQPLNSLDVPLKPRIPMETFRKVGIPPIIIALLSLASIIIVVLLIKVILDKYFL 60

Db 1 MESDSGLNNRDIIVPRKRRPQETPKKVGIPPIIAVLLSLIALVIVALLIKVILDKYF 60
Qy 61 LCGOPLHPIPRKQICDGLDCPLGEDBEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
Db 61 ICGSPLTFIQRGQLCDGHLDCASGEDBEHCVKDFPEKPGVAVRLSKDRSTLQVLDATGT 120
Qy 121 WFSACFNFTEALAEACRQWYSSKPTFRVIGIGDODLDVVEITENSQELMRNNSGP 180
Db 121 WASVCFNFTEALAKTACRQWYSSKPTFRVIGIGDODLDVVEITENSQELMRNNSGP 180
Qy 181 CLSGSLVSLHCLAGCKSLKTPRVVGGEEASVDSWPMQVSIQYDKQVCGSSILDEHWLIT 240
Db 181 CLSGSLVSLRCLDCGKSLKTPRVVGGVEAPVDSWPMQVSIQYDKQVCGSSILDEHWLIT 240
Qy 241 AAHCERKHTDVFHWKVRAGSDKLGSPSPSLAVAKIIIEFNPMYPKNDIALMKLOPFTF 300
Db 241 AAHCERKHTDVFHWKVRAGSDKLGSPSPSLAVAKIIIEFNPMYPKNDIALMKLOPFTF 300
Qy 301 SGTVRPCLPFFDEELTPATPLWIGWFTKONGKWSLILQASVQVLDSTRCNADAY 360
Db 301 SGTVRPCLPFFDEELTPATPLWIGWFTKONGKWSLILQASVQVLDSTRCNADAY 360
Qy 361 QGVVTEKMCAGIPEGVDTCCQDSGGPLMYQSDQVHVHVGIVSWGCGGPGSTPGVYTKV 420
Db 361 QGVVTEKMCAGIPEGVDTCCQDSGGPLMYQSDQVHVHVGIVSWGCGGPGSTPGVYTKV 420
Qy 421 SAYLNNIYNVWKAEL 435
Db 421 TAYLNNIYNVWKAEL 435

RESULT 3
Q4RR7; TETNG PRELIMINARY; PRT; 388 AA.
AC Q4RR7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 16 SCAP15002, whole genome shotgun sequence.
DE (Fragment)
GN ORFNames=GSTENG00030043001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Parra G., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Kellis M., Laffie J.N., Guigo R., Zody M.C., McKernan K.J., McEwan P., Bosak S.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
the early vertebrate proto-karyotype.",
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAB01015002; CAG08915.1; -; Genomic_DNA.

FT NON TER 1 388 388
SQ SEQUENCE 388 AA; 42696 MW; 07B0ADF4BA4C5736 CRC64;
Query Match 32.5%; Score 762; DB 2; Length 388;
Best Local Similarity 39.3%; Pred. No. 3.8e-56;
Matches 151; Conservative 56; Mismatches 121; Indels 56; Gaps 8;
Qy 95 PEGPAVAVRLSKDRSTLQVLDATGNSWFSACFNFTEALAEACRQWYSSKPTFRV-- 152
Db 13 PHSGSVFTRISPENSLLEIQGLKPTLWPCYERWNSSLGTLVCRQLGYLRLTKHGVL 72
Qy 153 -EIGDDODLDVVEIT-----ENSQELMRNNSGPGCLSGSLVSLHCLAGCKSLKTPRV 204
Db 73 TDIGENYTDGFIQITPEHNSLENIWQLR-----GSCVTGKVISLQCFEGCTRAKLPRII 127
Qy 205 GGEASVDSWPMQVSIQYDKQVCGSSILDPHWLTAACHCFRKHDT----- 250
Db 128 GGVETALGRWPQVSLYSSRHTCGGSIINSQWVTAACHCVKVASLVLGLSPNRLTS 187
Qy 251 -----VFNWVRA-----GSDKLGSPSPSLAVAKIIIEFNPMY---PKDN 287
Db 188 TSULLSEMSATGKLRVSSWVYAGIITRGSAKVAEHVGYAVEKII---YKNYHNRSHDG 244
Qy 288 DIALMKLQFPLTSGTVRPICLPFFDEELTPATPLWIGWFTKONGKWSLILQASVQ 347
Db 245 DIALKLRLTFLNFSDIIRPVCLPQYDYEPPGCTQCHWISGWYIQPEGVHSPDILKEAPVP 304
Qy 348 VIDSTRCNADAYQGVTEKMCAGIPEGVDTCCQDSGGPLMYQSDQ-VHVHVGIVSWGY 406
Db 305 IISTKRCNSCMYNGEITSRMLCAGYTEKGKACQDSGGPLVCQDENVWRLAGVVSWS 364
Qy 407 GCGGPGSTPGVYTKVASLNNIYNV 430
Db 365 GCAEPNHPGYTKVAEFLGWIDM 388
RESULT 4
TMPSD_MOUSE STANDARD; PRT; 543 AA.
AC QSU405; Q8CFE0; Q91VQ8;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
protease) (Membrane-type mosaic serine protease).
GN Name=Tmpsrsl3; Synonyms=Msp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC STRAIN=B5/EGFP, and FVB/N;
RC TISSUE=Mammary tumor, and Trophoblast stem cell;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan T.B., Moore T., Max S.I., Wang J., Hong L.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,


```

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RL the early vertebrate proto-karyotype.";
RN Nature 431:946-957(2004).
RP [2]
RG NUCLEOTIDE SEQUENCE.
RC Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAEE01014537; CAF97472.1; -; Genomic_DNA.
CC
FT NON_TER 1
FT NON_TER 359
SQ SEQUENCE 359 AA; 38943 MW; 4C14083C78233B37 CRC64;

Query Match 29.9%; Score 699.5; DB 2; Length 359;
Best Local Similarity 39.7%; Pred. No. 7.2e-51;
Matches 145; Conservative 57; Mismatches 138; Indels 25; Gaps 8;

QY 75 CDGELDCPLGEDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNWFSACDFNFTAL 134
DB 5 CDGVRDCTLGSDETAC-----VNLGND-NILQVKTSDQGRFLPVCYINGWDESLA 52

QY 135 ETACQMGVSSKPTTFAVEIGDPDQDLVVETENSQE---LRMR-NSSGPCILSGSLVSLH 190
DB 53 KETCTKLGRN---FVATNPSTQPKSPPTLTINSRSPYLQGRVTVVSSCPGQQTVALQ 109

QY 191 CLACGKSLKTRVVGGEASVDSPWQVSIQYDKOHVCGSILDPHWVLTAAHCFRKHDT 250
DB 110 CLDCQGRSTRIIGNVAKLQGWPMQMTLHFRGSHVCGGILISPDFVLTAAHCFPSNK 169

QY 251 VF---NWKVRAGSKLGSFPB-LAVAKIIIEFNPMYPKNDTALMKLQPLTFSGTVR 305
DB 170 LAILAENVEVSGVESLQKLPKPKYKILLSELNSTNDYDVALLEKLAAPVFDNVDQ 229

QY 306 PICLFFDEELTPATPLWIIWGFTKONGKRMDSILLQASVQVIDSTRCNADDAAYQGEVT 365
DB 230 PACLPSPDQILAPGTCQWTGTTGTTEDGSSSVSKSLMEVNIISDTVCNSVTVYKAVT 289

QY 366 EKMVCAGIPEGVDPCQDSGGLPMYO-SDQHWVVGIVSWYGGGSPGTPGVYTKVSAYL 424
DB 290 KNLKAGDLKGGKSCQDSGGLPVCQEDDRMYVYVGGITWSGGCQAKPGVTVYRVSVL 349

QY 425 NWIYN 429
DB 350 PWIYS 354

RESULT 6
ID TMSPD HUMAN STANDARD; PRT; 581 AA.
AC Q9BYE2; Q96YM4; Q96YJ8; Q9BYE1;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DE Transmembrane protease, serine 13 (EC 3.4.21.-) (Mosaic serine
DE protease) (Membrane-type mosaic serine protease).
GN Name=TMPRSS13; Synonyms=MSP, TMPRSS11;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1 AND 3), AND TISSUE SPECIFICITY.
RC TISSUE=Lung;
RX MEDLINE=21167393; PubMed=11267681; DOI=10.1016/S0167-4781(01)00184-1;
RA Kim D.R., Sharmir S., Inoue M., Kido H.;
RT "Cloning and expression of novel mosaic serine proteases with and
RT without a transmembrane domain from human lung.";
RL Biochim. Biophys. Acta 1518:204-209(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2).

```

```

RA Park T.J., Park W.J.;
RT "Homo sapiens transmembrane protease, serine 6 (TMPRSS6) mRNA.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 4).
RC TISSUE=Placenta; DOI=10.1038/ng1285;
RX PubMed=14702039;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togoya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukushima T.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC Human chromosome 11 international sequencing consortium;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -! SUBCELLULAR LOCATION: Type II membrane protein (potential).
CC -! ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=MSP, Large form, Membrane-type;
CC IsoId=Q9BYE2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9BYE2-2; Sequence=VSP_013103, VSP_013104;
CC Note=No experimental confirmation available;
CC Name=3; Synonyms=MSPS, Small form;
CC IsoId=Q9BYE2-3; Sequence=VSP_013099, VSP_013102;
CC Name=4;
CC IsoId=Q9BYE2-4; Sequence=VSP_013100, VSP_013101;
CC Note=No experimental confirmation available;
CC -! TISSUE SPECIFICITY: Isoform 1 and isoform 3 are predominantly
CC expressed in lung, placenta, pancreas, and prostate. Isoform 3 is
CC weakly expressed in testis and peripheral blood lymphocytes.
CC -! POLYMORPHISM: The repeat A-S-P-A-[GLQR] is polymorphic and the
CC number of copies varies between 12 to 14.
CC -! SIMILARITY: Belongs to the peptidase S1 family.
CC -! SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -! SIMILARITY: Contains 1 peptidase S1 domain.
CC -! SIMILARITY: Contains 1 SRCR domain.
CC -! CAUTION: Ref.2 has referred to this protein as TMPRSS6.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AB048796; BAB39741.1; -; mRNA.

```



```

Db 143 AQLGPPSYVSSDHLRVDALEBQFQDFVSNHLLSDDKVTALHHSVYMRG---CTSghv 199
Qy 187 VSLHCLACG-KSLKTPRVVGGEEASVDSPWQVSIQYDKQHVCGGSIIDPHWLVTAACHF 245
Db 200 VTLKCSACGTRGYSPRIVGGNMSSLTQWPQVSLQFQGYHLCCGSIITPLVITAAHCV 259
Qy 246 RKHTDVFNWVKRAGSDKL--GSFPLSLAVAKIIIEFNPMY-PK--DNDIALMKLOPPLTF 300
Db 260 YDLHPKSWTVQVGLVSLMDSVPVSHLVEKII---YHSKYKPKRLGNDIALMKLSEPLTF 316
Qy 301 SGTVPRIICLPFFDELTPTALWIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAY 360
Db 317 DETIQICLPNSEENFPDKLWTSWGAT-EDGDASPLVNAHAAVPLISNKNHRDYY 375
Qy 361 QGEVTEKMKCAGIPGEGVDTCCGSGGPLMYQSDQ-WHVVGIVSWGYGCGGSTPGVYTK 419
Db 376 GGIISPSMLCAGYLKGGVDCQDGGPLVCQERRLWKLVGATSPGIGCAEVNKPQVYTR 435
Qy 420 VSAYNLWYVNVKAEI 435
Db 436 ITSFLDWIHEQLERDL 451

RESULT 9
TPRS3_MOUSE STANDARD; PRT; 453 AA.
AC Q8KIT0; Q8VDE0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-MAY-2005 (Rel. 47, Last annotation update)
GN Name=tmpr33;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE, SUBCELLULAR LOCATION, AND FUNCTION IN ENAC
CLEAVAGE.
RX MEDLINE=22281255; PubMed=12393794; DOI=10.1093/hmg/11.23.2829;
RA Guipponi M., Vuagniaux G., Wattenhofer M., Shibuya K., Vazquez M.,
RA Dugherly L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,
RA Buchet K., Reymond A., Hummler E., Marzella P.L., Kudoh J.,
RA Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.;
RT "The transmembrane serine protease (TPRS3) mutated in deafness
RT DFNB8/10 activates the epithelial sodium channel (ENAC) in vitro.";
RL Hum. Mol. Genet. 11:2829-2836(2002).
CC -!- FUNCTION: Probable protease. Seems to be capable of activating
CC ENAC.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
CC reticulum.
CC -!- TISSUE SPECIFICITY: Expressed in the spiral ganglion, the cells
CC supporting the organ of Corti and the stria vascularis.
CC -!- PTM: Undergoes autoproteolytic activation.
CC -!- SIMILARITY: Belongs to the peptidase S1 family.
CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
CC -!- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AJ429216; CAD22137.1; -; Genomic_DNA.
CC EMBL; AJ300738; CAC83350.1; -; mRNA.
CC HSSP; P00760; 1EZX.
CC MEROPS; S01.079; -.
CC Ensembl; ENSMUSG00000024034; Mus musculus.
CC MGI; MGI:2155445; Tmprs3.
CC GO; GO:0016021; C:integral to membrane; TAS.

```

```

DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; Ldl_recept_a; 1.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PROSITE; PS01209; LDLRA_1; 1.
DR PROSITE; PS00068; LDLRA_2; 1.
DR PROSITE; PS00287; SRCR_2; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Endoplasmic reticulum; Glycoprotein; Hydrolase; Protease;
KW Serine protease; Signal-anchor; Transmembrane; zymogen.
FT TOPO_DOM 1 48
FT TRANSMEM 49 69
FT
FT TOPO_DOM 70 453
FT DOMAIN 72 108
FT DOMAIN 104 205
FT DOMAIN 217 448
FT ACT_SITE 257 257
FT ACT_SITE 304 304
FT ACT_SITE 400 400
FT SITE 216 217
FT CARBOHYD 221 221
FT DISULFID 73 85
FT DISULFID 79 98
FT DISULFID 92 107
FT DISULFID 129 194
FT DISULFID 142 204
FT DISULFID 207 324
FT DISULFID 242 258
FT DISULFID 338 406
FT DISULFID 369 385
FT DISULFID 396 424
FT CONFLICT 117 117
FT CONFLICT 246 246
SQ SEQUENCE 453 AA; 49492 MW; 1ABCBF10AF6E1EF6 CRC64;

Query Match 29.3%; Score 687; DB 1; Length 453;
Best Local Similarity 36.9%; Pred. No. 1.1e-49;
Matches 161; Conservative 72; Mismatches 153; Indels 50; Gaps 14;

Qy 22 IPMEFRKVGPIIIALLSLASIIIVVLIKVILD---KYFLCQGPLHFIPRKQLCDGE 78
Db 44 LPLKFPPIVIGIILALALA-----ICLGHFDCSGKY--RCHSSFKCIELTARCDGV 95
Qy 79 LDCPLGEBEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFTALAEATC 138
Db 96 SDCKNADEYRC-----VRVSGRAALQVFTAAA--WRTWCSDDMKSHYAKIAC 142
Qy 139 ROMGYSS-----KPTFRAVEIGPDOLDVVEITENSQELWRNSSGCPCLSGSL 186
Db 143 AQLGPPSYVSSDHLRVDALEBQFQDFVSNHLLSDDKVTALHHSVYMRG---CTSghv 199
Qy 187 VSLHCLACG-KSLKTPRVVGGEEASVDSPWQVSIQYDKQHVCGGSIIDPHWLVTAACHF 245
Db 200 VTLKCSACGTRGYSPRIVGGNMSSLTQWPQVSLQFQGYHLCCGSIITPLVITAAHCV 259
Qy 246 RKHTDVFNWVKRAGSDKL--GSFPLSLAVAKIIIEFNPMY-PK--DNDIALMKLOPPLTF 300
Db 260 YDLHPKSWTVQVGLVSLMDSVPVSHLVEKII---YHSKYKPKRLGNDIALMKLSEPLTF 316
Qy 301 SGTVPRIICLPFFDELTPTALWIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAY 360
Db 317 DETIQICLPNSEENFPDKLWTSWGAT-EDGDASPLVNAHAAVPLISNKNHRDYY 375
Qy 361 QGEVTEKMKCAGIPGEGVDTCCGSGGPLMYQSDQ-WHVVGIVSWGYGCGGSTPGVYTK 419
Db 376 GGIISPSMLCAGYLKGGVDCQDGGPLVCQERRLWKLVGATSPGIGCAEVNKPQVYTR 435

```

```
QY 420 VSAYLNWIYNVWMAEL 435
Db 436 ITSFLDWHQLEQLERDL 451

RESULT 10
Q9DGR2_XENLA
ID Q9DGR2_XENLA PRELIMINARY; PRT; 767 AA.
AC Q9DGR2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Embryonic serine protease-2.
GN Name-Xesp-2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20363741; PubMed=10903452; DOI=10.1016/S0378-1119(00)00225-0;
RA Yamada K., Takabatake T., Takeshima K.;
RT "Isolation and characterization of three novel serine protease genes
from Xenopus laevis."
RL Gene 252:209-216(2000).
DR EMBL; AB038497; BAB08217.1; -; mRNA.
DR HSP; P00760; IEZX.
DR MEROPS; S01.049; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0004263; F-1-chymotrypsin activity; IEA.
DR GO; GO:0008233; F-peptidase activity; IEA.
DR GO; GO:0005044; F-scavenger receptor activity; IEA.
DR GO; GO:0004295; F-trypsin activity; IEA.
DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR001254; Peptidase_S1_S6.
DR Pfam; PF00057; Srcr_receptor.
DR Kegg; K00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLA; 8.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS01209; LDLRA_1; 8.
DR PROSITE; PS00668; LDLRA_2; 2.
DR PROSITE; PS02040; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Protease.
SQ SEQUENCE 767 AA; 86001 MW; E0566A38796DE96E CRC64;

Query Match 29.3%; Score 686.5; DB 2; Length 767;
Best Local Similarity 38.3%; Pred. NO. 2.6e-49;
Matches 145; Conservative 58; Mismatches 153; Indels 23; Gaps 7;

QY 62 CQQLHFTIPRKQLCDGELDPLGEDEHCVKSPFEPGPAVAVRLSKDRSTLQVLDSATGNW 121
Db 395 CGSSVSCVLLSQWCDGVSDCPYGEDEMSCVLYPAD-----FQLQVTSVSAAW 443

QY 122 FSACFDNFTEALAEACRMQYSGKPTFRAVEI-----GPDDLDVVEITENSQELMRNS 177
Db 444 LPVCSDFWDDDFGRFACQDFGNGSSYNRTDLMSPYAPNGYFKLYSGYWRSKFYTSVQY 503

QY 178 SGPLSGSLSLHCLACGKSLKT--PRVVGEEASVDSPWQVSIQYDKQHCVCSSILDP 235
Db 504 SSYCVSGNVVSLHCISGVSNLSVSRIVGTTFANLGNWPNQVLYITGVLCGSGIISF 563

QY 236 HWVLTAAHC-FRKHTVFNKVRAGSKLGSFFSLA---VAKIIIEFNPMYPKNDIAL 291
Db 564 KWIIVTAHCYGVSYSSASGRVFAGTLPKTSYNNASAYFVERIIVHPGYKSTYNDIAL 623
```

```
QY 292 MKLQPLTFSGTVRPICLPFPDEBLTPATPLWIIGWFTKQNGKMSDILLOASVQVIDS 351
Db 624 MKLREIIFGTYTTPVCLPNSGMFWAEGATTTWISGWSTYE-GGSVSTYLQYAAIPLIDS 682

QY 352 TRCNADDAYQGEVTEKMMKAGIPGEGVDTCQDGGGGLPM-YQSDQWVHVGVISWGYCGG 410
Db 683 NVCMQSVYVNGQITSSMICAGYLSGGVDTCQDGGGGLPMKNGTWMVLVGTDSWGDGCAR 742

QY 411 PSTQGVYTKVSAYLNWIYN 429
Db 743 ANKPGVYGNVTTFLEWIS 761

RESULT 11
Q6NUFS_XENLA
ID Q6NUFS_XENLA PRELIMINARY; PRT; 722 AA.
AC Q6NUFS;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=Ovary;
RC MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=Ovary;
RC Klein S., Strausberg R.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC068636; AAH68636.1; -; mRNA.
DR HSSP; P00761; IAKS.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0004263; F-chymotrypsin activity; IEA.
DR GO; GO:0005044; F-scavenger receptor activity; IEA.
DR GO; GO:0004295; F-trypsin activity; IEA.
DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR001314; Peptidase_S1A.
```

```
DR InterPro; IPR001254; Peptidase_S1_S6.
DR InterPro; IPR001190; Ssr_receptor.
DR Pfam; PF00057; Ldl_recept_a; 4.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR PRINTS; PR00261; LDLRECEPTOR.
DR SMART; SM00192; LDLR; 7.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01209; LDLR1; 7.
DR PROSITE; PS50068; LDLR2; 1.
DR PROSITE; PS0287; SRCK2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR Hypothetical protein.
SQ SEQUENCE 722 AA; 80745 MW; 99D2B7391AE983FC CRC64;

Query Match 29.1%; Score 680.5; DB 2; Length 722;
Best Local Similarity 38.5%; Pred. No. 7.7e-49;
Matches 146; Conservative 57; Mismatches 153; Indels 23; Gaps 7;

QY 62 CGQPLHPIPRKQDGLDCLGDEEHCHVKSPFPGPAVAVRLSKDRSTLQVLDSATGNW 121
Db ||| : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : : |||
350 CGSSVSVLSSQWCDGSDCPYGEDMSC-----VRLGADFQLQVYITSVAW 398
QY 122 FSACDNFTALAEATACRMQYSSKPTFRAVEI---GPDQDLVDVEITENSQELMRNS 177
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
399 LPVCSYDNDPFRACQDFGYSYNYRDTLSPYAPNGYFKLYGWRSKFTYSVQY 458
QY 178 SGPLSGSLVSLHCLACGSKLKT--PRVVGGEASVDSPMQVSIQYDKQHVCGGSILDP 235
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
459 SSYCSYGNVSLHCLSCGYSVSVSRIVGTFANLGNWPMQVNIQYITVLGCGSIISP 518
QY 236 HWLTLAHC-FRKHTDVFNKVRAGSKDLGSPSLA---VAKIIIEFNPMPKNDIAL 291
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
519 KWIIVTAACHVYSGSASGRVFACTLTKPSYNASAYFVERIIVHPGYKSYTYNDIAL 578
QY 292 MKLQPLTFSGTVRPICLPFPFDEELTPATPLWIGWFTKQNGKMSDILLQASVOVIDS 351
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
579 MKLRDEITFGYTVQVCLPNSGMFWEAGTTWISGWSTYE--GGSVSTYLVQAAIPLIDS 637
QY 352 TRCNADDAAYQGEVTEKMKCAGIPEGGVDTCCQDSGGPLM-YOSDQHWVVGIVSWGYGCGG 410
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
638 NVNCNSVYVNGOITSSMICAGYLSGVDTCCQDSGGPLVKNKNGTWLWVGDTSWDGGCAR 697

QY 411 PSTPGYTKVSAYLWNIYN 429
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
698 ANKPGVYGNVTFLEWIYS 716

RESULT 12
QGSKTK7 HUMAN PRELIMINARY; PRT; 492 AA.
AC Q6GTK7.
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Transmembrane protease, serine 2 variant.
GN Name=TWPRSS2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
```

```
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Borak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Peripheral Nervous System;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Maruyama K., Sugano S.;
RT "Oligo-capping: a simple method to replace the cap structure of
RT eucaryotic mRNAs with oligoribonucleotides.";
RL Gene 138:171-174(1994).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Suzuki Y., Yoshitomo K., Maruyama K., Suyama A., Sugano S.;
RT "Construction and characterization of a full length-enriched and a 5'-
RT end-enriched cDNA library.";
RL Gene 200:149-156(1997).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Suzuki Y., Sugano S., Totoki Y., Toyoda A., Takeda T., Sakaki Y.,
RA Tanaka A., Yokoyama S.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC051839; AAHS1839.1; -; mRNA.
DR EMBL; AK222784; BAD96504.1; -; mRNA.
KW Protease; Transmembrane.
SQ SEQUENCE 492 AA; 53859 MW; C05B5531C8A311C7 CRC64;

Query Match 28.9%; Score 676.5; DB 2; Length 492;
Best Local Similarity 39.1%; Pred. No. 1e-48;
Matches 150; Conservative 57; Mismatches 128; Indels 49; Gaps 13;

QY 75 CDGELDCPLGDEEHCHVKSPFPGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNTEALA 134
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
133 CDGVSHCFGDEENRCVRLY--GP-----NFILQVYSQRKSWHFVQCDDWNYNG 181
QY 135 ETACRMQYSSKPTFRAVEIIGPDQDLVDVEITENSQELMRNSG-----PCL 182
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
182 RAACRDMGY--KNFYSSQ-----GIVDSGSGTSFMKLTNSAGNVDIYKLYHSDACS 232
QY 183 SGSLVSLHCLACGSKL---KTPRVVGGEASVDSPMQVSIQYDKQHVCGGSILDPHWL 239
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
233 SKAVVSLRACIACGVNLSSRSQRIVGGSALPGAWPMQVSLHVQNVHVGCGSIITPEWIV 292
QY 240 TAAHCFRKH-TDVFNVKVRAGSKLGSP-----PSLAVAKIIIEFNPMPY---PKNDIAL 291
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
293 TAAHCVEKPLNNPHWTAFAGILR-QSFMFYGAGYQVEKVI---SHPNVDSTKNDIAL 348
QY 292 MKLQPLTFSGTVRPICLPFPFDEELTPATPLWIGWFTKQNGKMSDILLQASVOVIDS 351
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
349 MKLQPLTFDLVLRVCLPFPNFMQLPQLCWIISGWGATEEK-GKTSVNLNAKVLIIET 407
QY 352 TRCNADDAAYQGEVTEKMKCAGIPEGGVDTCCQDSGGPLMYQSDQ-WHVVGIVSWGYGCGG 410
Db ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : ||| : : |||
408 QRCNSRYVDNLITPMTIAGFLQGNVDSQDGGSGPLVTSKXNIWLLIGTWSGSGCAK 467
```


QY 411 PSTPGVYTKVSAYNLWYNNWKA 434
 Db 468 AYRPGVIGNVWFTDIWYQMRAD 491

RESULT 13
 TMPS2_HUMAN
 ID TMPS2_HUMAN STANDARD; PRT; 492 AA.
 AC O15393; Q9EXX1;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Transmembrane protease, serine 2 precursor [EC 3.4.21.-] [Contains:
 DE Transmembrane protease, serine 2 non-catalytic chain; Transmembrane
 DE protease, serine 2 catalytic chain].
 GN Name=TMPS2; Synonyms=PRSS10;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97468144; PubMed=9325052; DOI=10.1006/geno.1997.4845;
 RA Paoloni-Giacobino A., Chen H., Peitsch M.C., Rossier C.,
 RA Antonarakis S.E.;
 RT "Cloning of the TMPS2 gene, which encodes a novel serine protease
 RT with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3.";
 RL Genomics 44:309-320(1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21309069; PubMed=11414763; DOI=10.1006/geno.2001.6551;
 RA Teng D.-H., Chen Y., Lian L., Ha P.C., Tavtigian S.V., Wong A.K.C.;
 RT "Mutation analyses of 268 candidate genes in human tumor cell lines.";
 RL Genomics 74:352-364(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE, AND MUTAGENESIS.
 RX MEDLINE=21139112; PubMed=11245484;
 RA Afar D.E.H., Vivanco I., Hubert R.S., Kuo J., Chen E., Saffran D.C.,
 RA Raitano A.B., Jakobovits A.;
 RT "Catalytic cleavage of the androgen-regulated TMPS2 protease results
 RT in its secretion by prostate and prostate cancer epithelia.";
 RL Cancer Res. 61:1686-1692(2001).
 RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=21104370; PubMed=11169526;
 RA Vaarala M.H., Porvari K.S., Kellokumpu S., Kyllonen A.P., Vihko P.T.;
 RT "Expression of transmembrane serine protease TMPS2 in mouse and
 RT human tissues.";
 RL J. Pathol. 193:134-140(2001).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Activated by
 CC cleavage and secreted.
 CC -!- TISSUE SPECIFICITY: Expressed strongly in small intestine. Also
 CC expressed in prostate, colon, stomach and salivary gland.
 CC -!- SIMILARITY: Belongs to the peptidase S1 family.
 CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -!- SIMILARITY: Contains 1 LDL-receptor class A domain.
 CC -!- SIMILARITY: Contains 1 peptidase S1 domain.
 CC -!- SIMILARITY: Contains 1 SRCR domain.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

 CC EMBL; U75329; AAC51784.1; -; mRNA.
 CC EMBL; AF123453; AAD37117.1; -; mRNA.
 CC EMBL; AF270487; AAK29280.1; -; mRNA.
 CC HSP; P00760; 1EZx.
 CC MEROPS; S01.247; -;
 CC Ensembl; ENSG00000184012; Homo sapiens.
 CC HGNC; HGNC:11876; TMPS2.

DR MIM; 602060; -.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0008236; F:serine-type peptidase activity; TAS.
 DR InterPro; IPR001254; LDL receptor A.
 DR InterPro; IPR001254; Peptidase_S1_S6.
 DR InterPro; IPR001314; Peptidase_S1A.
 DR InterPro; IPR001190; Srcr_receptor.
 DR Pfam; PF00089; Trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR SMART; SM00192; LDLA; 1.
 DR SMART; SM00202; SR; 1.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01209; LDLRA_1; 1.
 DR PROSITE; PS00668; LDLRA_2; 1.
 DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
 DR PROSITE; PS0287; SRCR_2; 1.
 DR PROSITE; PS0240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;
 KW Signal-anchor; Transmembrane; Zymogen.
 FT CHAIN 1 255
 FT Transmembrane protease, serine 2 non-
 FT catalytic chain.
 FT CHAIN 254 492
 FT Transmembrane protease, serine 2
 FT catalytic chain.
 FT TOPO_DOM 1 84
 FT TRANSMEM 85 105
 FT Signal-anchor for type II membrane
 FT protein (Potential).
 FT FT Extracellular (Potential).
 FT FT LDL-receptor class A.
 FT FT SRCR.
 FT FT Peptidase S1.
 FT FT Charge relay system (By similarity).
 FT FT Charge relay system (By similarity).
 FT FT Charge relay system.
 FT FT Cleavage (Potential).
 FT FT N-linked (GlcNAc. .) (Potential).
 FT FT N-linked (GlcNAc. .) (Potential).
 FT FT By similarity.
 FT FT By similarity.
 FT FT By similarity.
 FT FT By similarity.
 FT FT By similarity.
 FT FT By similarity.
 FT FT Interchain (between non-catalytic and
 FT catalytic chains) (By similarity).
 FT FT By similarity.
 FT FT By similarity.
 FT FT By similarity.
 FT FT K -> N (in dbSNP:1056602).
 FT FT /FTID=VAR_011692.
 FT FT R->Q: Loss of cleavage.
 FT FT S->A: Loss of activity.
 FT FT M -> V (in Ref. 3).
 FT FT I -> L (in Ref. 1).
 FT FT E -> Q (in Ref. 1).
 FT FT RAD -> KAN (in Ref. 1).
 FT FT 491
 FT SS 53891 MW; CAB44FD174A9076B CRC64;
 SQ SEQUENCE 492 AA; 53891 MW; 53891 MW; 53891 MW;

Query Match 28.8%; Score 673.5; DB 1; Length 492;
 Best Local Similarity 38.8%; Pred. No. 1.8e-48;
 Matches 149; Conservative 58; Mismatches 128; Indels 49; Gaps 13;
 QY 75 CDGELDCPLGEDEBHCVKSPFPGPAVAVRLSKDRSTLQVLDSATGNWFSACFDNFTEALA 134
 Db 133 CDGVSHCPGDEENCRVLY--GP-----NFIQMYSSQKSWHPVCQDDWNNYVG 181
 QY 135 ETACRMQYSGKPTFAVEIGDPDQDLVVEITENQELMRNMSG-----PCL 182
 Db 182 RAACRDMGY--KXNFYSSQ-----GIVDDSGSTFMKLNTSAGNVDYTKLYHSDACS 232
 QY 183 SGLSVSLHCLACKSL---KTPRVVGGREASVDSWPQVSIQYDKQHCGGSIILDPHWL 239
 Db 233 SKAVVSLRCIACGVNLNLSRQSRIVGGESALPGANFWQVSLHVQNVHVCSSIIITPEWIV 292

RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment.";
RL Genome Res. 13:2285-2270 (2003).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM E).
RX PubMed=14702039; DOI=10.1038/ngl285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahori K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamiyama M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa K., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Saeki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45 (2004).
RN [5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS A AND E).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [6]
RP SUBCELLULAR LOCATION AND FUNCTION IN ENAC CLEAVAGE.
RX MEDLINE=22281255; PubMed=12393794; DOI=10.1093/hmg/11.23.2829;
RA Guipponi M., Vucnetiaux G., Wattenhofer M., Shibuya K., Vazquez M.,
RA Doughterty L., Scamuffa N., Guida E., Okui M., Rossier C., Hancock M.,
RA Buchet K., Raymond A., Hummel E., Marzella P.L., Kudoh J.,
RA Shimizu N., Scott H.S., Antonarakis S.E., Rossier B.C.;
RT "The transmembrane serine protease (TMPRSS3) mutated in deafness
RT DFNB8/10 activates the epithelial sodium channel (ENaC) in vitro.";
RL Hum. Mol. Genet. 11:2829-2836 (2002).
RN [7]
RP VARIANTS DFNB8/DFNB10 CYS-251 AND LEU-404.

RE MEDLINE=21354482; PubMed=11462234; DOI=10.1002/humu.1159;
RA Masmoudi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
RA Pappaavas M.P., Driira M., Elgaied-Boulila A., Wattenhofer M.,
RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
RT "Novel missense mutations of TMPRSS3 in two consanguineous Tunisian
RT families with non-syndromic autosomal recessive deafness.";
RL Hum. Mutat. 18:101-108 (2001).
RN [8]
RP VARIANTS DFNB8/DFNB10 TRP-109; PHE-194 AND ARG-407, AND VARIANTS
RP ILR-53; SER-111 AND VAL-253.
RX MEDLINE=21317610; PubMed=11424922; DOI=10.1136/jmg.38.6.396;
RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
RA Kudoh J., Shibuya K., Antonarakis S.E., Bonne-Tamir B.,
RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
RT "Novel mutations of TMPRSS3 in four DFNB8/10 families segregating
RT congenital autosomal recessive deafness.";
RL J. Med. Genet. 38:396-400 (2001).
RN [9]
RP VARIANT DFNB8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426
RX MEDLINE=21904597; PubMed=11907649; DOI=10.1007/s00109-001-0310-6;
RA Wattenhofer M., Di Iorio V., Rabionet R., Dougherty L., Pampalos A.,
RA Schwede T., Montserrat-Sentis B., Arbones L., Iliades T.,
RA Paquidibiscaglia A., D'Amelio M., Alwan S., Rossier C., Dahl H.-H.M.,
RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
RA Antonarakis S.E.;
RT "Mutations in the TMPRSS3 gene are a rare cause of childhood
RT nonsyndromic deafness in Caucasian patients.";
RL J. Mol. Med. 80:124-131 (2002).
RN [10]
CC -1- FUNCTION: Probable protease. Seems to be capable of activating
CC ENaC.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
CC reticulum.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=5;
CC Name=A;
CC IsoId=P57727-1; Sequence=VSP_005391;
CC Name=B; Synonyms=C;
CC IsoId=P57727-2; Sequence=VSP_005391;
CC Name=D;
CC IsoId=P57727-3; Sequence=VSP_005392;
CC Name=E; Synonyms=Truncated, TADG-12V;
CC IsoId=P57727-4; Sequence=VSP_005393, VSP_005394;
CC Name=E;
CC IsoId=P57727-5; Sequence=VSP_013184;
CC -1- TISSUE SPECIFICITY: Expressed in many tissues including fetal
CC cochlea. Isoform T is found at increased levels in some
CC carcinomas.
CC -1- PTM: Undergoes autolytic activation.
CC -1- DISEASE: Defects in TMPRSS3 are a cause of childhood-onset
CC autosomal recessive neurosensory deafness 8 (DFNB8) [MIM:601072].
CC -1- DISEASE: Defects in TMPRSS3 are a cause of congenital autosomal
CC recessive neurosensory deafness 10 (DFNB10) [MIM:605316].
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
CC -1- SIMILARITY: Contains 1 LDL-receptor class A domain.
CC -1- SIMILARITY: Contains 1 peptidase S1 domain.
CC -1- SIMILARITY: Contains 1 SRCR domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AF201380; AAC37012.1; -; mRNA.
CC EMBL: AB038157; BAB20077.1; -; mRNA.
CC EMBL: AB038158; BAB20078.1; -; mRNA.
CC EMBL: AB038159; BAB20079.1; -; mRNA.
CC EMBL: AB038160; BAB20080.1; -; mRNA.
CC EMBL: AK172842; BAB1806.1; -; mRNA.
CC EMBL: AY359458; AAC08823.1; -; mRNA.
CC EMBL: BC074846; AAH74846.1; -; mRNA.
CC EMBL: BC074847; AAH74847.1; -; mRNA.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 19:48:25 ; Search time 5647 Seconds
(without alignments)
4378.768 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPSPDQLNSLDVVKPLRP.....VTKVSALYLNWYVWKAEL 435

Scoring table:

	BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0		Ygapext 0.5	
Fgapop 6.0		Fgapext 7.0	
Delop 6.0		Delext 7.0	

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10803530/runat.05122005.083215.2119/app.query.fasta_1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530 @CGN 1.1 7415 @runat.05122005.083215.2119 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
GenEmbl.*
1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.hcg.*
15: gb.pl.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	1305	6	AX076190 Sequence
2	2338	99.8	2038	6	ARI142620 Sequence
3	2338	99.8	2038	6	BD137129 Human pro

	4	2338	99.8	2038	6	AR5637419	Sequence
5	2337	99.8	1314	6	CQ924452	Sequence	
6	2337	99.8	2021	8	AK172766	Homo sapi	
7	2337	99.8	2079	8	BC011703	Homo sapi	
8	2337	99.8	2081	6	AR454680	Sequence	
9	2337	99.8	2081	6	AF179224	Homo sapi	
10	2335	99.7	2121	6	AR559487	Sequence	
11	2333	99.6	2590	6	CQ869992	Sequence	
12	2329	99.4	1479	6	AX076192	Sequence	
13	2328	99.4	2627	6	CQ869994	Sequence	
14	2324	99.2	2070	6	BD235831	A novel m	
15	2324	99.2	2079	6	AR232520	Sequence	
16	2324	99.2	2079	6	AR454679	Sequence	
17	2324	99.2	2079	6	AX207967	Sequence	
18	2324	99.2	2079	8	AF216312	Homo sapi	
19	2319	99.0	2137	6	AX207899	Sequence	
20	2297.5	98.1	2063	6	CQ875679	Sequence	
21	2297.5	98.1	2063	6	CS112264	Sequence	
22	2297.5	98.1	2063	6	AX092380	Sequence	
23	2297.5	98.1	2063	6	AX376262	Sequence	
24	2297.5	98.1	2063	6	AX395214	Sequence	
25	2297.5	98.1	2063	6	AX697206	Sequence	
26	2297.5	98.1	2063	8	AY358530	Homo sapi	
27	2167	92.5	1784	8	BC012752	Homo sapi	
28	2044	87.3	1973	8	BC004855	Homo sapi	
29	1844	78.7	2060	9	BC021368	Mus muscu	
30	1844	78.7	2268	9	AY043240	Mus muscu	
31	1782	75.1	2229	6	CQ869989	Sequence	
32	1546.5	66.0	955	6	CQ729939	Sequence	
33	1287	55.0	699	6	EL13203	Human cDNA	
34	1259	53.8	1189	6	AR559493	Sequence	
35	860	36.7	2947	5	CR855611	Xenopus t	
36	806	34.4	1008	6	BD194909	86 human	
37	806	34.4	1008	6	CQ855237	Sequence	
38	704.5	30.1	2992	9	BC042878	Mus muscu	
39	699.5	29.9	2067	6	CS075313	Sequence	
40	699.5	29.9	2067	6	AX149581	Sequence	
41	699.5	29.9	3320	9	BC085323	Mus muscu	
42	698.5	29.8	1434	6	AR583622	Sequence	
43	697.5	29.8	1314	6	AX190543	Sequence	
44	697.5	29.8	1314	6	AX190546	Sequence	
45	697.5	29.8	1689	6	CS075405	Sequence	

ALIGNMENTS

RESULT 1
AX076190
LOCUS AX076190 1305 bp linear PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104141.
ACCESSION AX076190
VERSION AX076190.1 GI:12710815
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suendermann,B., Hofmann,U., Matzku,S. and Wilbert,O.
TITLE Seripancrin
JOURNAL Patent: WO 0104141-A 1 18-JAN-2001;
MERCK PATENT GmbH (DE)
FEATURES
source
1..1305
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..>1305
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28479.1"
/db_xref="GI:12710816"


```
Qy 41 LeuAlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db 320 CTGGCGAGTATCATATTGTTCTCTCATCAAGGTGATTCTCGATAAATACTACTTC 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTCTCGGGGAGCCCTCCACCTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGAC 439
Qy 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db 440 TGTCTCTTGGGGAGGAGCAGGAGCACTGTGTCAAGAGCTTCCCCCAAGGGCTGCACTG 499
Qy 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 500 GCAGTCCGCCCTCCCAAGACCGATCCACACTGCAGAGTCTCGCTGACAGAGCCCTGTAGGCAG 559
Qy 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 560 TGGTTCTCTGCTGTTTTCGACAACTTCACAGAAGCTCTCGCTGACAGAGCCCTGTAGGCAG 619
Qy 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCCAAGCAGGATCTG 679
Qy 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 680 GATGTTGTTGAATCACAAGAAACAGCAGGAGCTTCGATCGCGAATCAAGTGGGGCCC 739
Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 740 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACCTGCTTGTGCTGTGGGAGAGCCCTGAAGACC 799
Qy 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 800 CCCCCTGTGTGGTGGGAGGAGGCCCTCTGTGGATTCTTGGCCCTTGGCAGGTTCAGCATC 859
Qy 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 860 CAGTACGACAAACACGACGCTGTGGAGGAGCATCCTGGACCCCACTGGGTCTCTCAGC 919
Qy 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 920 GCAGCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAAGTGGGGCAGGCTCA 979
Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 980 GACAAACTGGGCAGCTTCCATCCCTGGCTGTGGCCAGAGCATCATCAATCAATTTCAAC 1039
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1040 CCCATGTATCCCCAAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCACTACTTC 1099
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCAGCCACC 1159
Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 1160 CACATCTGGATCATTTGGATGGGGCTTTAGCAGCAGATGGAGGAGAGATGTCTGACATA 1219
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1220 CTGCTGCAGCGCTCAGTCCAGGTATTGACACACACGCGTCAATGTCAGACAGATCGGTAC 1279
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1280 CAGGGGGAGTCAACGAGAAGATGATGTGTGAGGAGCATCCCGAAGGGGTGTGGACACC 1339
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1340 TGCCAGGGGTGACAGTGGTGGGGCCCTGTGATCAATCTGACCAAGTGGCATGTGTGGTGGGC 1399
```

```
Qy 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1400 ATCGTTAGCTGGGCTATGGCTGGGGGGCCGAGCACCCAGGAGTATACACCAAGGTC 1459
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1460 TCAGCTATCTCAACTGGATCTACAATGTCTGGAAAGGCTGAGCTG 1504
RESULT 3
LOCUS BD137129 Human protease molecule. 2038 bp DNA linear PAT 18-SEP-2002
DEFINITION BD137129
ACCESSION BD137129
VERSION BD137129.1 GI:23232074
KEYWORDS JP 2002508970-A/6.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2038)
AUTHORS Bandman,O., Hillman,J.L., Yue,H., Guegler,K.J., Corley,N.C.,
Tang,T.Y. and Shah,P.
Human protease molecule
Patent: JP 2002508970-A 6 26-MAR-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002508970-A/6
PD 26-MAR-2002
PF 12-JAN-1999 JP 2000540252
PI 16-JAN-1998 US 09/008271
PI OLGA BANDMAN,JENNIFER L HILLMAN,HENRY YUE,KARL J GUEGLER,NEIL
C CORLEY.
PI TOM Y TANG,PURVI SHAH
PC C12N15/09,A61K38/46,C07K16/40,C12N1/19,C12N1/21,C12N5/10 PC
,C12N9/48,C12N9/64,
PC C12Q1/68,C12N15/00,A61K37/54,C12N5/00
CC 1337018
FH Key Location/Qualifiers
FT source 1..2038
FT Location/Qualifiers
1..2038
/organism='Homo sapiens (human)'
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores: 1.6e-207 Length: 2038
Pred. No.: 2338.00 Matches: 434
Score: 2338.00 Conservative: 1
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 99.77% Indels: 0
Query Match: 99.83% Gaps: 0
DB: 6
US-10-803-530-2 (1-435) x BD137129 (1-2038)
Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 200 ATGGATCTTGACATGATCACTCTGAACAGCCTCGATGTCAAACCCCTGGCAACCC 259
Qy 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSer 40
Db 260 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCCATCATCATAGCACTACTGAGC 319
Qy 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db 320 CTGGCGAGTATCATATTGTTCTCTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379
Qy 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 380 CTCTCGGGGAGCCCTCCACCTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGAC 439
```



```
QY      81  CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGAGGACGAGAGACACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499

QY      101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCCGCTCTCCAGAGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAAC 559

QY      121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db      560 TGGTTCTCTGCTGTTTCGACAACTTCACAGAGCTCTCGCTGAGACAGCTGTAGGCAG 619

QY      141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db      620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACGAGTCTG 679

QY      161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTTGAAATCACAGAAACACAGCAGAGCTTCGCATGCGGAACTCAAGTGGGCC 739

QY      181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGTTCTCCCTGCACTGTCTTGGCTGTGGGAGAGCTGAGACC 799

QY      201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      800 CCCCCTGTGTGGTGGGGAGAGGCGCTCTGTGGATTCTTGGCTTGGCAGGTCAGCATC 859

QY      221 GlnTyrAspLysGlnHisValCysGlyCysIleSerIleLeuAspProHisTrpValLeuThr 240
Db      860 CAGTACGACAAACAGCAGCTGTGGAGGAGCATCTCGACCCCACTGGGTCCTCACG 919

QY      241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      920 GCAGCCCACTGCTTCAGAAACATACCGATGTGTTCACCTGGAAGGTGCGGGCAGGCTCA 979

QY      261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      980 GACAAACTGGGCAGCTTCCCATCCCTGGCTGTGCCAAGATCATCATTTGAATTCAAC 1039

QY      281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db      1040 CCCATGTATCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCATTTTC 1099

QY      301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db      1100 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGGCCACC 1159

QY      321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIle 340
Db      1160 CCACTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGAAGATGCTGTGACATA 1219

QY      341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db      1220 CTGCTGCAGGCGTCAGTTCAGGTCAATGACAGCACACGTTGCAATGCAGACGATGCGTAC 1279

QY      361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db      1280 CAGGGGAGNATCACCGAGAAGATGATGTGCAGGCATCCCGAAGGGGGTGTGGACACC 1339

QY      381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db      1340 TGCCAGGGGTGACAGTGGTGGGCCCTGTATGTACCAATCTGACCAAGTGGCATGTGGTGGGC 1399

QY      401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db      1400 ATCGTAGCTGGGCTATGGCTGCGGGGGCCGAGCACCCAGGAGTATATACCAAGGTC 1459

QY      421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1460 TCAGCCTATCTCACTGGATCTCAATGTCTGGGAAGGCTGAGCTG 1504
```

```
RESULT 4
AR637419
LOCUS      AR637419          2038 bp          DNA          linear          PAT 20-APR-2005
DEFINITION Sequence 18 from patent US 6855811.
ACCESSION AR637419
VERSION    AR637419.1  GI:627771139
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 2038)
AUTHORS    Bandman,O., Hillman,J.L., Yue,H., Guegler,K.J., Corley,N.C.,
            Tang,Y.T. and Shah,P.
TITLE      Human protease molecules
JOURNAL    Patent: US 6855811-A 18 15-FEB-2005;
            Incyte Pharmaceuticals, Inc.; Palo Alto, CA
FEATURES   Location/Qualifiers
            source
            1..2038
            /organism="unknown"
            /mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.:      1.6e-207      Length:      2038
Score:          2338.00      Matches:    434
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.77%      Mismatches: 0
Query Match:    99.83%      Indels:     0
DB:             6            Gaps:       0
US-10-803-530-2 (1-435) x AR637419 (1-2038)
QY      1  MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db      200 ATGGATCTCTGACAGTAGTATCAACTCTGACAGAGCTTGATGTCAAAACCCCTGCGCAAAACC 259

QY      21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      260 CGTATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATGACTACTGAGC 319

QY      41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db      320 CTGGCAGGATATCATATTGTTGTTGCTCATCAGGTGATTCTGGATAAATACTACTTC 379

QY      61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db      380 CTCTGGGGCAGGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGAGAGCTGGAC 439

QY      81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 499

QY      101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCGCTCTTCCAAGGACCGATCCACACTGCAGGTGCTGAGCTCGGCCACAGGGAAC 559

QY      121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db      560 TGGTTCTCTGCTGTTTCGACAACTTCAGAGAGCTCTCGCTGAGACAGCTGTAGGCAG 619

QY      141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db      620 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACGAGTCTG 679

QY      161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTTGAAATCACAGAAACACAGCAGAGCTTCGCATGCGGAACTCAAGTGGGCC 739

QY      181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGTTCTCCCTGCACTGTCTTGGCTTGGCAGGTCAGCATC 799

QY      201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
```



```
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 970 CTCTGGATCATTCGATGGGCTTTTACGAAGCAGATGGAGGAGAGATGCTCAGACATCTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTCAGCGGTCAGTCAGGTCATTTGACAGCACACGCGTCAATGCAGACCATCGTACCAG 1089
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1090 GGGGAAGTCACCGAAGATGATGTGCGAGCATCCCGAAGGGGTGTGACACCTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGGTGACAGTGTGGGCGCTGATGTACCAATCTGACCACTGCGCATGTGTGGGCATC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrIysValSer 421
Db 1210 GTTAGCTGGGCTATGGCTGGCGGGCGCGAGCACCCAGAGGATATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1311

RESULT 6
AK172766
LOCUS
DEFINITION
Homo sapiens cDNA FLJ23927 fis, clone COL05059, highly similar to
Transmembrane protease, serine 4 (EC 3.4.21.-).
ACCESSION
AK172766
VERSION
AK172766.1 GI:47077742
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominae; Homo.
REFERENCE
1
AUTHORS
Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
UNPUBLISHED
2 (bases 1 to 2021)
Sugano,S. and Suzuki,Y.
Direct Submission
Submitted (22-APR-2004) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
(E-mail:flcdna@mail.ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction,
5'- & 3'-end one pass sequencing; Department of Virology and Human
Genome Center, Institute of Medical Science, University of Tokyo
(partly supported by Science and Technology Agency).
FEATURES
Location/Qualifiers
1..2021
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL05059"
/tissue_type="colon"
/clone_lib="COL"
/note="cloning vector pME18SFL3"
148..1461
/note="unnamed protein product"
/codon_start=1
/protein_id="BAD18749.1"
/db_xref="GI:47077743"
/translation="MLQDPSDQPLNSLDVKPLRKPRIPMETFRKVGIPITIALISLA
SIIVVLIKVIDKYFLCGQLHFIPRKQLCDGLDCLPGLGDEHCKVSPFEGPAV
```

ORIGIN

```
Alignment Scores:
Pred. No.: 1,96e-207 Length: 2021
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 8 Gaps: 0

US-10-803-530-2 (1-435) x AK172766 (1-2021)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 157 GATCCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGGCAACCCCGT 216
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 217 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCATCTACTAGCCTG 276
QY 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 277 GCAGATATCATCATTTGTTGTTGCTCATCAAGGTGATTTCTGGATAAATACTACTTCTCTC 336
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 337 TGGCGGACGCTTCCACTTCATCCGAGAGAGAGCTGTGTGACGAGAGCTGGACTGT 396
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 397 CCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTTCAGTGGCA 456
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 457 GTCGCGCTCTCAAGAGACCGATCCACACTGCGAGGTGTGTGACTCGGCCACAGGGAAC 516
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 517 TTCTCTGCTGTTTCGACAACTTCAGAGAGCTTCGCTGAGACAGCTGTAGGCAGATG 576
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 577 GGTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCACAGCAGATCTGGAT 636
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 637 GTTGTGAAATCACAGAAAAACAGCCAGAGCTTCGATTCGGAATCTCAAGTGGGCCCTGT 696
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 697 CTCTCAGGCTCCCTGCTCTCCCTGCACCTGCTTGCCTGTGGAGAGCTTGAAGACCCCC 756
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 757 CGTGTGGTGGGGGAGGAGGCTCTGTGTGATTTCTTGCCCTTGGCAGGTCAAGTCCAG 816
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 817 TACGACAAACAGCAGCTCTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCTACGGCA 876
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 877 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGCGAGCTCAGAC 936
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 937 AAACCTGGGCAGCTTCCCATCTCCCTGGCTGTGCCAAGATCATCATTAATTAACCCC 996
```

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 997 ATGTACCCCAAGACAATGACATGCCCTCATGAAGTGCAGTTCCTCCACTTCTCA 1056

QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 1057 GGCACAGTCAGGCCCATCTGCTGCCCTCTTTGATGAGAGCTCCTCCAGCCACCCA 1116

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 DB 1117 CTCTGGATCTGGATGGGCTTTACGAAGCAGAAATGAGGGAAGATGCTGACATCTG 1176

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
 DB 1177 CTCGAGGCGTCAGTCAGGTCATTGACAGCACACCGTGCATGCGAGAGTCGTCACG 1236

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 DB 1237 GGGGAAGTCACCGAAGATGATGTGCGAGCATCCCGAAGGGGGTGTGGACACCTGC 1296

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
 DB 1297 CAGGCTGACAGTGGTGGGCCCTGATGTACCAATCTGACAGTGCATGTGGTGGCATC 1356

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 DB 1357 GTTAGTGGGCTATGGCTGCGGGGGCCGAGCACCAGGAGTATACCAAGGCTCTCA 1416

QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
 DB 1417 GCCTATCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1458

RESULT 7

BC011703

LOCUS BC011703 2079 bp mRNA linear PRI 08-MAR-2005
 DEFINITION Homo sapiens transmembrane protease, serine 4, transcript variant 1, mRNA (CDNA clone MGC:19490 IMAGE:3610695), complete cds.

ACCESSION

BC011703

VERSION

BC011703.2

GI:33991397

KEYWORDS

MGC.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 2079)

REFERENCE

AUTHORS

Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Ustin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullaly SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodriguez S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green EP, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalek U, Smalios DE, Schnurch A, Schein JE, Jones SJ and Marra MA. Mammalian Gene Collection Program Team

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

CONSRTM

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (30-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.mci.nih.gov>

On Aug 20, 2003 this sequence version replaced gi:15079794.

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Diatchenko, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, O.L., Magliello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Tourgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 27 Row: f Column: 19

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 34304348.

FEATURES

Location/Qualifiers

1..2079

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:19490 IMAGE:3610695"

/tissue type="Pancreas, adenocarcinoma"

/clone_lib="NIH MGC 39"

/lab_host="DH10B-R"

/note="Vector: pOTB7"

1..2079

/gene="TMPRSS4"

/note="synonyms: TMPRSS3, MT-SP2"

/db_xref="GeneID:56649"

/db_xref="MIM:606565"

208..1521

/gene="TMPRSS4"

/codon_start=1

/product="transmembrane protease, serine 4, isoform 1"

/protein_id="AAH11703.1"

/db_xref="GI:15079795"

/db_xref="GeneID:56649"

/translation="MLQDPDSQPLNSLDVRLKPRIPMETFRKVGIPILIALISLA

SIIIVVLIKVILOKYLFCGLPHFIPRKOLCDGELDCPLGDEDEHCVKSPPEGPV

AVRLSKRSTLQVLSATGNWFSACFDNFTALACROMGYSSKPTFRVVEIGPDQ

DLDDVTEITNSQELRMNNSGFLSGSLVSLHCLAGSLKTPRVVGGEEASVDSWP

QVSYQDKHQVGGSLDPLHWLTAHCRKHTDVFNMKVRAGSKLGSFPPSLAVAKI

IIEFNPKYNDIALLMKLQFPLTFSGTVPICLPFFDELTPTPLWIIQWGTQK

NGKMSDILLQASVQVIDSTRCNADDAYQGEVTERMCAGIEGGVDTCQSGSGFLM

YQSDQHWVGVIVSGVCGGPGSTPGVYTKVSAYLNIWYVWKAEL"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

2..2079

2337.00

100.00%

99.79%

8

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

2079

434

0

0

0

0

US-10-803-530-2 (1-435) x BC011703 (1-2079)

QY

2

AspProAspSerAspGlnProLeuAsnSerLeuAspVallyProLeuArglyProArg 21

DB

217

GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAACCCCGT 276

QY 22 IleProMetGluThrPheArgGlyValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 277 ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTAGACCTG 336
QY 42 AlaSerIleIleValValValValValValValValValValValValValValValVal 61
Db 337 GCGAGTATCATATGTGGTGTCTCATCAAGGTGATTCGTGATAAATACTACTTCTCTC 396
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 397 TCGGGGAGGCTCTCCACTTCATCCCGAGGAGCAGCTGTGTGACGAGAGCTGGAGCTGT 456
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 457 CCCTTGGGGAGGAGCAGGAGGAGCTGTGTCAAGAGCTTCCTCCGAGGCGCTGAGTGGCA 516
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyLeuTrp 121
Db 517 GTCCGCTCTCCAGGACCGATCCACACTGCAGGTGTGAGCTCGGCCACAGGGAAGTGG 576
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 577 TTCTCTGCTGTTCGACAACTTCACAGAAAGCTTCGCTGACAGAGCTGTAGGAGATG 636
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 637 GGTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGGCCACAGACCGAGATCTGGAT 696
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerClyProCys 181
Db 697 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGATCGGGAATCAAGTGGGCCCTGT 756
QY 182 LeuSerClySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 757 CTCTAGGCTCCCTGTGCTCCCTGCACTGTCTGTGCTGTGGAGAGCTTGAAGACCCC 816
QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 817 CGTGTGGTGGTGGGAGAGGCGCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAG 876
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 877 TAGCAAAACAGCAGCTGTGTGGAGGAGCATCTGGACCCCACTGGTCTCTCAGCGCA 936
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 937 GCCCACTGCTTCAGAAACATACCGATGTGTCAACTGGAAGGTGCGGCGAGGCTCAGAC 996
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleGluPheAsnPro 281
Db 997 AAACCTGGCAGCTTCCCATCCCTGGCTGTGGCCAGATCATCATTTGAATTCACCCC 1056
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1057 ATGATACCCCAAGACAAATGACATGCCCTCATGAAGTGCAGTTCCCACTTCTCTCA 1116
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1117 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGTCACTCCAGGACAGCCCA 1176
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1177 CTCTGGATCATTTGGTGGGCTTTACGAGCAGAAATGGAGGAGATGTCTGACATACTG 1236
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaThrGln 361
Db 1237 CTGACGGCGTCAGTCCAGGTCATTGACAGACACGGTGCATTCAGACAGATGGGTACCAG 1296
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1297 GGGGAAGTCACCGAGAAGATGATGTGTGACAGGATCCCGAAGGGGGTGTGGACACCTGC 1356

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1357 CAGGGTGACAGTGGTGGGCCCTGTGATGTACCAATCTGACCAGTGGCATGTGTGGGCATC 1416
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1417 GTTAGCTGGGCTATGGCTGCGGGGGCCGAGCAGCCAGAGGATATACCAAGGTCTCA 1476
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 GCCTATCTCAACTGGATCTACAAATGTCTGGAAGGCTGAGCTG 1518
RESULT 8
LOCUS AR454680
DEFINITION Sequence 7 from patent US 6682890.
ACCESSION AR454680
VERSION AR454680.1 GI:42688149
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2081)
AUTHORS Mack,D., Gish,K.C. and Wilson,K.E.
TITLE Methods of diagnosing and determining prognosis of colorectal cancer
JOURNAL Patent: US 6682890-A 7 27-JAN-2004;
Protein Design Labs, Inc.; Fremont, CA
FEATURES
source Location/Qualifiers
1..2081
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,04e-207 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0
US-10-803-530-2 (1-435) x AR454680 (1-2081)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 224 GATCTGACAGTGATCAACCTCTGAACAGCTCGATGTCAACCCCTGCGCAACCCCGT 283
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 284 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 343
QY 42 AlaSerIleIleValValValValValValValValValValValValValValValVal 61
Db 344 GCGAGTATCATATGTGGTGTCTCATCAAGGTGATTCGTGATAAATACTACTTCTCTC 403
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 404 TCGGGGAGGCTCTCCACTTCATCCCGAGGAGCAGCTGTGTGACGAGAGCTGAGACTGT 463
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 464 CCCTTGGGGAGGAGCAGGAGCAGCTGTGTCAAGAGCTTCCCGAAGGCGCTGAGTGGCA 523
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 524 GTCCGCTCTTCCAGAGCCGATCCACACTGCAGGTGTGTGGACTCGGCCACAGGGAATGG 583
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 584 TTCTCTGCTCTTTCGACAACTTCACAGAAAGCTCTCGCTGAGACAGGCTGTAGCAGATG 643
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161

Db 644 GCGTACAGCAGCAACCCACTTTCAGAGCTGTGGAGATTGGCCCAAGACAGGATCTGGAT 703
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 704 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCGCAACTCAAGTGGCCCTGT 763
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlySerLeuHisThrPro 201
Db 764 CTCTCAGGCTCCCTGGTCTCCCTGCACATGTCTTGTGCTGTGGGAAGAGCTGAAGACCC 823
Qy 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 824 CGTGTGTGGTGGGAGAGAGCCCTCTGTGGATTCTTGGCCCTTGCGAGGTTCAGCATCCAG 883
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 884 TACGACAAACAGCAGCTGTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCTACGGCA 943
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 944 GCCCACTGCTTCAGGAACATACCATGTGTTCAACTGGAAGGTGCGGCGAGCTCAGAC 1003
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 1004 AAATCTGGCAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGAATTCAACCCC 1063
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1064 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTCACTTCTCA 1123
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1124 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGTATGAGGAGCTCACTCCAGCCACCCA 1183
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1184 CTCTGGATCATTTGGATGGGCTTTACGAAGCAAGATGGAGGAAAGATGTGACATACTG 1243
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaValTrpGln 361
Db 1244 CTGCAGGCGTCAGTCCAGGTCAATGCAGCACACGGTGCATGCAGAGCATGCGTACCAG 1303
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1304 GGGGAAGTCACCCAGAGATGATGTGTGAGGATCCCGAAGGGGTGTGGACACCTGC 1363
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1364 CAGGCTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACCACTGTGGTGGGCATC 1423
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1424 GTTACTGGGCTATGGCTGGCGGGCCCGAGCACCCAGGAGTATACCAAGGTCTCA 1483
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1484 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1525

RESULT 9

AF179224
LOCUS AF179224 2081 bp mRNA linear PRI 08-JUN-2000
DEFINITION Homo sapiens transmembrane serine protease 3 (TMPRSS3) mRNA,
complete cds.

ACCESSION AF179224

VERSION AF179224.1 GI:8347148

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominoidea; Homo.

REFERENCE 1 (bases 1 to 2081)

AUTHORS

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 2081

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q23.3"

/tissue_type="pancreatic carcinoma"

1. 2081

/gene="TMPRSS3"

215..1528

/gene="TMPRSS3"

/codon_start=1

/product="transmembrane serine protease 3"

/protein_id="AAF74526.1"

/db_xref="GI:8347149"

/translation="MLQPDSDQPLNSLDVKLRPRIPMETFRKVGIPPIIALLSLA

SIIVLVILKVLKYLFLCQPLHFI PRKOLCDGELDCPLGEDHEHCVKSPSPGPVAV

AVRLKORSTLQVLDATGNWFSACFDNFTALAEATACROMGYSSKPTFRVAEIGPDQ

DLVVVITENSQVELRMNNSGFCLSGLVSLHCLACGKSLKTPRVVGGEEASVDSWPW

QVSIQDKHQVCGGSLIDPHWLVTAAHCFRKHTDVFNVKVRAGSDKLGSPFSLAVAKI

IIIFNPYPKNDIALMKLQPLTFSGTVRPICLPFFDEELTPATPLWIIWGGFTKQ

NGKMSDILLQASQVIDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCQSDSGGFLM

YQSDQHVVGIVSWGCGGPGSPGTGVTYKVSAYLNLWVYVWKAEI"

215..307

/gene="TMPRSS3"

/note="Region: cytoplasmic domain"

308..373

/gene="TMPRSS3"

/note="Region: transmembrane domain"

374..1525

/gene="TMPRSS3"

/note="Region: extracellular domain"

602..604

/gene="TMPRSS3"

/note="glycosylation site"

746..748

/gene="TMPRSS3"

/note="glycosylation site"

order(800..802,1142..1144)

/gene="TMPRSS3"

/note="disulfide bond"

order(902..904,950..952)

/gene="TMPRSS3"

/note="disulfide bond"

947..949

/gene="TMPRSS3"

/note="active site"

1082..1084

/gene="TMPRSS3"

/note="active site"

order(1280..1282,1328..1330)

/gene="TMPRSS3"

/note="disulfide bond"

1373..1375

/gene="TMPRSS3"

/note="active site"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

2.04e-207

2337.00

Length:

2081

Matches:

434

Wallrapp,C., Hahnel,S., Muller-Pillasch,F., Burghardt,B.,
Iwamura,T., Ruthenburger,M., Lerch,M.M., Adler,G. and Gress,T.M.
A novel transmembrane serine protease (TMPRSS3) overexpressed in
pancreatic cancer
Cancer Res. 60 (10), 2602-2606 (2000)
10825129
2 (bases 1 to 2081)
Wallrapp,C. and Gress,T.M.
Direct Submission
Submitted (20-AUG-1999) Internal Medicine I, University of Ulm,
Robert-Koch-Street 8, Ulm, Baden-Wuerttemberg 89081, Germany

Percent Similarity: 100.00%		Conservative: 0
Best Local Similarity: 100.00%		Mismatches: 0
Query Match: 99.79%		Indels: 0
DB: 8		Gaps: 0
US-10-803-530-2 (1-435) x AF179224 (1-2081)		
QY	2	AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
Db	224	GATCTGACAGTGATCAACCTCTGAACAGCTTCGATGTCAAACCCCTGGCAAAACCCCGT 283
QY	22	IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db	284	ATCCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCTG 343
QY	42	AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrrPheLeu 61
Db	344	GGAGTATCATCATTTGGTGTCTCATCAAGGTGATTTCTGGATAAATACTACTTCCTC 403
QY	62	CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db	404	TGCGGGAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGACTGT 463
QY	82	ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db	464	CCCTTGGGGAGGAGCAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCCTGCAGTGGA 523
QY	102	ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db	524	GTCCGCCCTCTCCAGAGCCGATCCACTGCTGAGGTGCTGGAGCTCGGCCACAGGAACTGG 583
QY	122	PheSerAlaCysPheAspAsnPhetrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db	584	TTCTCTGCTGTTTCGAACTTCACAGAGCTCTCGCTGAGACAGCTGTAGCAGATG 643
QY	142	GlyTrpSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db	644	GGCTACAGCAGAAACCCACTTTCAGAGCTGTGAGATTTGGCCAGACACAGGATCTGGAT 703
QY	162	ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db	704	GTGTGTGAATCACAGAAACAGCCAGGAGTTTCGATCGGAACTCAAGTGGGCCCTGT 763
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db	764	CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCTGTGGAGAGAGCTGAACACCCC 823
QY	202	ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db	824	CGTGTGTGGTGGGAGAGGCTCTGTGGATTTCTTGGCCTTGGCAGGTGAGATCCAG 883
QY	222	TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db	884	TACGACAAACAGCAGCTGTGTGGAGGAGCATCTGGACCCCCACTGGTCCCTCACGGCA 943
QY	242	AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db	944	CCCCACTGCTTCAGAAACATACCGATGTGTCAACTTGAAGGTGGCGGAGGCTCAGAC 1003
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db	1004	AACTGGGAGCTTCCCATCCCTGGCTGTGGCCAGATCATCATTTGAATTCACCCCC 1063
QY	282	MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db	1064	ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGAGTTCCTCACTTCTCA 1123
QY	302	GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
Db	1124	GGCAGACTGAGCCCATCTGTCTGCCCTTCTTGATGAGGAGCTCACTCCAGCCACCCCA 1183
QY	322	LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
Db	1184	CTCTGGATCATTTGATGGGGCTTTTACGAAGCAGAAATGAGGAGATGTCTGACATACTG 1243
QY	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrrGln 361
Db	1244	CTGCAGGCGTCAGTCCAGGTCAATTGACAGCACACGGTGCATTCGAGACGATGGTACCAG 1303
QY	362	GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db	1304	GGGGAAGTCAACGAGAAGATGATGTGTGACGAGCATCCCGGAAGGGGTGTGGACACCTGC 1363
QY	382	GlnGlyAspSerGlyGlyProLeuMetTyrrGlnSerAspGlnTrpHisValValGlyIle 401
Db	1364	CAGGTGACAGTGGTGGGCCCTTGATGTACCAATCTGACCACTGGCATGTGGTGGGCATC 1423
QY	402	ValSerTrpGlyTyrrGlyCysGlyIleProSerThrProGlyValTyrrLysValSer 421
Db	1424	GTTAGCTGGGCTATGGCTGCGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTCA 1483
QY	422	AlaTyrrLeuAsnTrpIleTyrrAsnValTrpLysAlaGluLeu 435
Db	1484	GCCTATCTCAACTGGATCTACAAATGTGTGGAAGGCTGAGCTG 1525
RESULT 10		
AR559487		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
FEATURES		
source		
1..2121		
/organism="unknown"		
/mol_type="genomic DNA"		
ORIGIN		
Alignment Scores:		
Pred. No.: 3-2e-207		
Score: 2335.00		
Percent Similarity: 99.77%		
Best Local Similarity: 99.77%		
Query Match: 99.70%		
DB: 6		
US-10-803-530-2 (1-435) x AR559487 (1-2121)		
QY	1	MetAspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysPro 20
Db	277	ATGGATCTCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGGCAAAACC 336
QY	21	ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db	337	CGTATCCCCATGGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGCACTACTGAGC 396
QY	41	LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrrPhe 60
Db	397	CTGGCGAGTATCATCATTTGGTGTCTCTCATCAAGGTGATTTCTGATAAATACTACTTC 456
QY	61	LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db	457	CTCTGGGCGACCTCTCCACTTTCATCCGAGGAGAGCAGCTGTGTGACGGAGAGCTGGAC 516
QY	81	CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db	517	TGTCCCTTGGGGAGGAGCAGGAGGAGCTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTG 576


```
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 577 GCAGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGCTGGACTCGGCACAGGGAAC 636
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 637 TGGTCTCTGCTGTTTCGACAACTTCAGAAAGCTCTCGCTGAGACAGCCTGTAGGAG 696
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 697 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCAGACCAGGATCTG 756
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 757 GATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCC 816
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 817 TGTCTCTCAGGCTCCCTGCTCCCTGCACCTGCTGTGGCTGTGGAAAGAGCCTGAAGACC 876
QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 877 CCCCCTGTGGTGGGGAGAGGCCCTCTGTGGATTCTTGCCCTTGGCAGGTGAGCATC 936
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db 937 CAGTACGACAAACAGACACGCTCTGTGGAGGAGCATCTCGACCCCACTGGGTCTCAGC 996
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db 997 GCAGCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGGGGCAGGCTCA 1056
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db 1057 GACAACTGGGAGCTTCCATCCCTGGCTGTGGCCAGATCATCATCAATGCAATCAAC 1116
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1117 CCCATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTTCCTCACTTTC 1176
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 1177 TCAGGCACAGTCAGGCCCATCTGTCTGCCCTCTTTTGTATGAGGAGCTCACTCCAGCCAC 1236
QY 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIle 340
Db 1237 CCACTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGGAGGAAAGATGTCTGACATA 1296
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1297 CTGCTGCAGCGCTCAGTCCAGGTCAATTGACAGCACACGCGTGCAATGCAGACGATCGTAC 1356
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThr 380
Db 1357 CTGGGGGAAGTCAACGAGAAGATGATGTGTGAGGATCCCGAAAGGGGTGTGACAC 1416
QY 381 CysGlnGlyAspSerGlyLysProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1417 TGCCAGGGTGACAGTGGTGGGCCCTTGATGTACCAATCTGACAGTGGCATGTGTGGGGC 1476
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db 1477 ATCGTTAGTGGGCTATGGCTCGGGGGCCCGAGCACCCAGGGGTATACACCAAGTTC 1536
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1537 TCAGCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1581
```

RESULT 11

```
CQ869992
LOCUS 2590 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 413 from Patent WO2004074320.
ACCESSION CQ869992
```

```
VERSION CQ869992.1 GI:51999709
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Morris,D.W., Morris,D.W. and Malandro,M.S.
TITLE Novel therapeutic targets in cancer
JOURNAL Patent: WO 2004074320-A 413 02-SEP-2004;
Sages Discovery, Inc. (US)
FEATURES
source
1..2590
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6,32e-207 Length: 2590
Score: 2333.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.62% Indels: 0
Dbs: 6 Gaps: 0
US-10-803-530-2 (1-435) x CQ869992 (1-2590)
QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db 236 ATGGATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGGCAAAACC 285
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db 286 CGTATCCCATGGAGACCTTCAGAAAGGTGGGGATCCCATCATCATAGCACTACTGAGC 345
QY 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
Db 346 CTGGCGAGTATCATATTGTGTTGCTCTCATCAAGGTGATTCTGGATAAATACTACTTTC 405
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db 406 CTCTCGGGCAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGAGAGCTGGAC 465
QY 81 CysProLeuGlyGluAspGluLysCysValLysSerPheProGluGlyProAlaVal 100
Db 466 TGTCTCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGCGCTGCAGTG 525
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db 526 GCAGTCCGCTCTCCAAAGACCGATCCACACTGCGAGGTGCTGGACTCGGCCACAGGGAAC 585
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 586 TGGTCTCTGCTGTTTCGACAACTTCAGAAAGCTCTCGCTGAGACAGCCTGTAGGACG 645
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db 646 ATGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCCAAGCATCTG 705
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db 706 GATGTTGTTGAAATCACAGAAACAGCCAGGAGCTTCGCATCGGAACTCAAGTGGGCC 765
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db 766 TGTCTCTCAGGCTCCCTGCTCTCCCTGCACCTGTCTTGTCTGTGGGAAGAGCCTGAAGAC 825
QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db 826 CCCCCTGTGGTGGGTGTGGAGGAGGCTCTGTGGATTCTTGGCCTTGGCAGGTTCAGCATC 885
```

QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
|
|
|
Db 886 CAGTACGACAAACAGCACGTCTGTGGAGGAGCATCTCGGACCCCATCGGTCTCAGC 945
|
|
|
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTyrLysValArgAlaGlySer 260
|
|
|
Db 946 GAGGCCCATCTGCTTCAGGAAACATACCGATGTGTTCACCTGGAAGGTGCGGCGAGCTCA 1005
|
|
|
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleIleIleIle 280
|
|
|
Db 1006 GACAACTGGGAGCTTCCATCTCCCTGGCTGTGGCCAGATCATCATTTGAATCAAC 1065
|
|
|
QY 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
|
|
|
Db 1066 CCAATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCATTTCCACTCATTTTC 1125
|
|
|
QY 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
|
|
|
Db 1126 TCAGGCACAGTCAGGCCCATCTGTCTGGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1185
|
|
|
QY 321 ProLeuTyrIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
|
|
|
Db 1186 CCACCTCTGATCATTTGGATGGGGCTTTACGAAGCAGAATGGAGGGAAGATGTCTGACATA 1245
|
|
|
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
|
|
|
Db 1246 CTGCTCGAGGCTCAGTCCAGTCAATGACAGCACACCGTGCAATGCAGAGATGCGGTAC 1305
|
|
|
QY 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
|
|
|
Db 1306 CAGGGGGAAGTCACCCAGAGAAGATGATGTGTCAGGATCCCGAAGGGGTGTGGACACC 1365
|
|
|
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheIleValGly 400
|
|
|
Db 1366 TCCAGGGTGACAGTGGTGGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGGC 1425
|
|
|
QY 401 IleValSerTyrGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
|
|
|
Db 1426 ATCGTTAGTGGGGCTATGGCTGCGGGGGCCGAGCACCCAGGAGTATACCAAGGTC 1485
|
|
|
QY 421 SerAlaTyrLeuAsnTyrIleTyrAsnValTyrLysAlaGluLeu 435
|
|
|
Db 1486 TCAGCCTATCAACTGGATCTACAAATGTCTGGAAGGTGAGCTG 1530
|
|
|

RESULT 12
AX076192
LOCUS AX076192
DEFINITION Sequence 3 from Patent WO0104141.
ACCESSION AX076192
VERSION AX076192.1 GI:12710817
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
Suendermann, B., Hofmann, U., Matzku, S. and Wilbert, O.
Seripancrin
Patent: WO 0104141-A 3 18-JAN-2001;
MERCK PATENT GmbH (DE)
Location/Qualifiers
1..1479
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1479
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28480.1"
/db_xref="GI:12710818"
/translation="MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPPIIALLSLASI
IIVVVLKILDKYFLCGQPLHPIPRKQLCDGELDCLGDEBEHCVKSPFEGPAVAV

CDS

source

FEATURES

ORGANISM

KEYWORDS

VERSION

ACCESSION

LOCUS

DEFINITION

SEQUENCE

KEYWORDS

SOURCE

ORGANISM

FEATURES

ORIGIN

Alignment Scores:
Pred. No.: 7.38e-207 Length: 1479
Score: 2329.00 Matches: 432
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.44% Indels: 0
DB: 6 Gaps: 0
US-10-803-530-2 (1-435) x AX076192 (1-1479)
QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
|
|
|
Db 1 ATGGATCTGACAGTGTATCAACCTCTGACAGCCTCGATGTCAAAACCCCTCGCAAAACC 60
|
|
|
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
|
|
|
Db 61 CGTATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGC 120
|
|
|
QY 41 LeuAlaSerIleIleIleValValIleValIleLysValIleLeuAspLysTyrTyrPhe 60
|
|
|
Db 121 CTGGCAGTATCATCATTTGGTGTCTCTCATCAAGGTGATCTGGATAATACTACTATTC 180
|
|
|
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
|
|
|
Db 181 CTCTGCGGGAGGCTCTCCACTTCATCCCGAGGAGCAGCTGTGTGACGAGAGCTGGAC 240
|
|
|
QY 81 CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
|
|
|
Db 241 TGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCCCTGCAAGT 300
|
|
|
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
|
|
|
Db 301 GCAGTCGCGCTCTCCAAGACCGATCCACACTCGAGGTGTGGACTCGGACACAGGGAAC 360
|
|
|
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
|
|
|
Db 361 TGGTCTCTGCGCTGTTTCGACAACTTCACAGAAAGCTCTCGCTGAGACAGCTGTAGGCAG 420
|
|
|
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
|
|
|
Db 421 ATGGGCTACAGCAGCAAAACCACTTTTCAGAGCTGTGGAGATTGGCCAGACAGGATCTG 480
|
|
|
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
|
|
|
Db 481 GATGTGTTGAAATCATCAGAAAACAGCCAGGAGCTTCGATCGGAACTCAAGTGGGCC 540
|
|
|
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
|
|
|
Db 541 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACATGTCTTGGCTGTGGGAAGAGCTTGAAGACC 600
|
|
|
QY 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTyrProTyrGlnValSerIle 220
|
|
|
Db 601 CCCCCTGTGTGGTGGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTCAGCATC 660
|
|
|
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
|
|
|
Db 661 CAGTACGACAAACAGCACGTCTGTGGAGGAGGATCTCTGGATTTCTTGGCTTGGCAGGTCAG 720
|
|
|
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTyrLysValArgAlaGlySer 260
|
|
|
Db 721 GCAGCCCACTGCTTTCAGGAAACATACCGATGTCTTCAACTGGAAGGTGCGGCGAGGCTCA 780
|
|
|
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsn 280
|
|
|

RLSKDRSTLOVDSATCNWFSACFDNFTEALAEATACRMGYSSKPTTRAVEIGPDQDL
DVVEITENSBELNRNNSGCLGSVSLHCLACGKSLKTPRVVGGSEASVDSWPQV
SIQDKOHHVCGGSLIDPHWLVTAHCFKHTDVENWKRAGSDKLGFFPSLAIAKII
IEFNPMYKNDNDIALMKLPFLPSTGRPICLPFFDELTTPATPLMIIGWFTKQNG
KQSDIILLOQASVQIDSTRNADDAVQCEVTEKMCACGIPGGVDTCQGSGLPMVQ
SDOWHVVGIVSWGVCGGPSTPGVYTKVSAYLNIYNWKDRTIQRSCNSPGTGLVIQ
QPAVPLMELASRHTHLTTTCVCFSLFPALQPLSLCPFPWD"

Db 781 GACAACTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCAC 840
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 841 CCCATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTCACTTTC 900
Qy 301 SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db 901 TCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTTTGATGAGAGCTCACTCCAGCCACC 960
Qy 321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 961 CCACCTCTGGATCATTTGGATGGGGCTTACGAAGCAGAAATGGAGGGAAGATGCTGCACATA 1020
Qy 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr 360
Db 1021 CTGTCGAGCGCTCAGTCAGGTCAATTCACAGCACACGCGTGCATTCGACAGCATCGGTAC 1080
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1081 CAGGGGGAAGTCAACGAGAAGATGATGTGCAGGCATCCCGGAAGGGGGTGTGGACACC 1140
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db 1141 TGCCAGGGTGACAGTGGTGGGCCCTGATGACCAATCTGACCACTGGCATGTGTGGGCG 1200
Qy 401 IleValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysVal 420
Db 1201 ATCGTTAGCTGGGGCTATGGCTGCGGGGCGCGAGCACCCAGGAGTATACACCAAGGTC 1260
Qy 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLys 432
Db 1261 TCAGCCTATCTCACTGGATCTACAATGTCTGGAAG 1296

RESULT 13
LOCUS CQ869994 2627 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 415 from Patent WO2004074320.
ACCESSION CQ869994
VERSION CQ869994.1 GI:51999710
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Morris,D.W., Morris,D.W. and Malandro,M.S.
TITLE Novel therapeutic targets in cancer
JOURNAL Patent: WO 2004074320-A 415 02-SEP-2004;
Sagres Discovery, Inc. (US)

FEATURES
source
1..2627
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1.88e-206 Length: 2627
Score: 2328.00 Matches: 433
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.40% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x CQ869994 (1-2627)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 266 GATCTCGACAGTGATCAACTCTGACACAGCTCGATGTCAAAACCCCTGCGCAACCCCGT 325
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSerLeu 41

Db 326 ATCCCCATGGAGAGCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTTGACGCTG 385
Qy 42 AlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 386 CGAGATATCATCATTTGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTC 445
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 446 TCGGGGAGCGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGACTGT 505
Qy 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 506 CCTTGGGGGAGGAGGAGACACTGTGTCAAGAGCTTCCCGAAGGGCCCTGCACTGGCA 565
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 566 GTCCGCTCTCCAAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAACCTG 625
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 626 TTCTCTGCTCTTTCGACAACTTCACGAAGCTCTCGCTGAGACAGCCCTGTAGGAGATG 685
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 686 GGCTACAGCAGCAAAACCCACTTTCAGACTGTGGAGATTGGCCCAAGACAGGATCTGGAT 745
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 746 GTTGTGTAATCAGCAAAAACAGCAGGAGCTTCGCATCGGNACTCAAGTGGGCGCTGT 805
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 806 CTCTCAGGCTCCCTGGTCTCCCTGCTGCTGCTTGGCTGTGGGAAGAGCTGAAGACCCCC 865
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGluValSerIleGln 221
Db 866 CGTGTGGTGGGTGGAGGAGCGCTCTGTGATCTTGGCCCTTGGCAGTTCAGCATCCAG 925
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 926 TACGACAAAACAGCAGCTCTGTGGAGGAGCATCTGGACCCCCACTGGTCTCTCAGGCA 985
Qy 242 AlahisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 986 GCCCACTCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGAC 1045
Qy 262 LysLeuGlySerPhePheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 1046 AAACCTGGGCAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1105
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1106 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTCACTTTCTCA 1165
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1166 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 1225
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1226 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACATACTG 1285
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1286 CTGCAGGGCTCAGTCCAGGTCAATTGACAGCACACCGTCAATGCGAGAGCATGCGTACCAG 1345
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1346 GGGGAAGTCAACGAGAGATGATGTGTCAGGCATCTCCGGAAGGGGGTGTGGACACCTGC 1405
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401

QY 421 rAlaTyrLeuLeuThrPheTyrAspValTrpIysAlaGluLeu 435
Db 1483 AGCCTATCTCAACTGGATCTACAATGTCTGGAAAGGCTGAGCTG 1525

RESULT 15
AR232520
LOCUS AR232520 2079 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6455668.
ACCESSION AR232520
VERSION AR232520.1 GI:27274657
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2079)
AUTHORS Mack,D., Gish,K.C. and Wilson,K.E.
TITLE Methods of diagnosing colorectal cancer, compositions, and methods
of screening for colorectal cancer modulators
JOURNAL Patent: US 6455668-A 1 24-SEP-2002;
Eos Biotechnology, Inc.; South San Francisco, CA
FEATURES
Location/Qualifiers
1..2079
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.3e-206 Length: 2079
Score: 2324.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.23% Indels: 1
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x AR232520 (1-2079)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCTGACAGTATCAACCTCTGACAGCCTCGATGTCAAAACCCCTGCGCAAAACCCCGT 276

QY 22 IleProMetGluThrPheArgLys-ValGlyIleProIleIleAlaLeuLeuSerIle 41
Db 277 ATCCCATGGAGACCTTCAGAAAGTGTGGGATCCCATCATCATAGCACTACTGAGCCT 336

QY 41 uAlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheIle 61
Db 337 GCGGAGTATCATATTGTGTGTTCTCATCAAGTGATTCTGGATAAATACTACTTCTCT 396

QY 61 uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db 397 CTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTG 456

QY 81 eProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db 457 TCCCTTTGGGGAGAGCAGAGAGACCTGTGTCAAGAGCTTCCCGAAGGGCCCTGCGATGGC 516

QY 101 aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db 517 AGTCGCCCTCTCCAAGGACCGATCCACACTGCGAGGTGTGGACTCGGCCACAGGAACCTG 576

QY 121 pPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMe 141
Db 577 GTTCTCTGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGACCTGTAGGCAGAT 636

QY 141 tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db 637 GGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACCTGGA 696

QY 161 pValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCy 181
Db 697 TGTGTGTGAATACACAGAAAACAGCCAGGAGCTTCGATGCGGAACCTCAAGTGGGCGCTG 756

QY 181 sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201

Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTGCTGTGGGAAGACGCTGAAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGl 221
Db 817 CCGTGTGTGGTGGGAGGAGGCTCTGTGGATTTCTTGGCTTGGCAGGTGAGCATCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACAGCACGCTCTGTGGAGGAGCATCTCTGACCCCACTGGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValAAlaGlySerAs 261
Db 937 AGCCCATCTGCTTCAGAAACATACCGATGTGTCAACTGGAAAGTGGCGGAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCATCCCTGGCTGTGGCCAAGATCATCATTAATTAATCAAGCCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTTCTTC 1116
QY 301 rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 1117 AGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleIle 341
Db 1177 ACTCTGGATCATTTGGATGGGCTTTTACGAAGCAGAAATGGAGGAAGATGTCTGACATACT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGl 361
Db 1237 GCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCATATGCAGACGATGCCATCCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 1297 GGGGGAAGTCAACGAGAAGATGATGTGTGAGGCATCCCGAAGGGGGTGTGGACACCTG 1356
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI 401
Db 1357 CCAGGGTCACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSe 421
Db 1417 CGTTAGCTGGGCTATGGCTGCGGGGCCCGAGCACCCCGAGAGTATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1519

Search completed: December 6, 2005, 21:56:59

Job time : 5677 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:53:15 ; Search time 166 Seconds
(without alignments)
1094.914 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVKPLRPK.....VTKVSYALNINYNWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA Main:*

- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	435	US-10-803-530-2	Sequence 2, Appli
2	2342	100.0	435	US-10-473-127-1578	Sequence 1578, Ap
3	2342	100.0	435	US-10-473-127-1597	Sequence 1597, Ap
4	2338	99.8	435	US-09-968-415-6	Sequence 6, Appli
5	2338	99.8	435	US-10-180-719-6	Sequence 6, Appli
6	2338	99.8	435	US-10-473-127-1596	Sequence 1596, Ap
7	2338	99.8	435	US-11-045-577-6	Sequence 6, Appli
8	2337	99.8	437	US-09-851-588-8	Sequence 8, Appli
9	2337	99.8	437	US-10-295-027-779	Sequence 779, App
10	2337	99.8	437	US-10-295-027-791	Sequence 791, App
11	2337	99.8	437	US-10-295-027-831	Sequence 831, App
12	2337	99.8	437	US-10-295-027-1156	Sequence 1156, Ap
13	2337	99.8	437	US-10-173-999-89	Sequence 89, Appli
14	2337	99.8	437	US-10-473-127-1581	Sequence 1581, Ap
15	2337	99.8	437	US-10-473-127-1586	Sequence 1586, Ap
16	2337	99.8	437	US-10-473-127-1601	Sequence 1601, Ap
17	2337	99.8	437	US-10-473-127-1602	Sequence 1602, Ap
18	2337	99.8	437	US-10-994-117-7	Sequence 7, Appli
19	2337	99.8	437	US-10-991-287-7	Sequence 7, Appli
20	2339	99.4	492	US-10-473-127-1579	Sequence 1579, Ap
21	2339	99.4	492	US-10-473-127-1598	Sequence 1598, Ap
22	2319	99.0	437	US-09-776-191-4	Sequence 4, Appli
23	2319	99.0	437	US-10-156-214A-4	Sequence 4, Appli
24	2319	99.0	437	US-10-473-127-1583	Sequence 1583, Ap
25	2319	99.0	437	US-10-473-127-1588	Sequence 1588, Ap
26	2297.5	98.1	432	US-09-888-257A-7	Sequence 7, Appli
27	2297.5	98.1	432	US-09-946-374-275	Sequence 275, App

ALIGNMENTS

RESULT 1

US-10-803-530-2
; Sequence, 2, Application US/10803530
; Publication No. US20040146981A1
; GENERAL INFORMATION:
; APPLICANT: Dartow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/10/803,530
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/09/607,745
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-803-530-2

Query Match	100.0%;	Score 2342;	DB 4;	Length 435;
Best Local Similarity	100.0%;	Pred. No. 1.9e-219;	Mismatches 0;	Indels 0; Gaps 0;
Matches 435;	Conservative 0;			
Qy	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPITIIALLSLASIIIVVVLKIVLDKYYF	60	
Db	1	MDPDSQPLNSLDVKPLRPKPRIPMETFRKVGIPITIIALLSLASIIIVVVLKIVLDKYYF	60	
Qy	61	LCGQPLHPIPRKQICDGLDCPLGDEBHCVKSPFEGPAVRLSKDRSTLOVLSATGN	120	
Db	61	LCGQPLHPIPRKQICDGLDCPLGDEBHCVKSPFEGPAVRLSKDRSTLOVLSATGN	120	
Qy	121	WFSACDFNFTALASTACRMGYSSKPTFRAVEIGPDODLDVVEITENSQELMRNSSGP	180	
Db	121	WFSACDFNFTALASTACRMGYSSKPTFRAVEIGPDODLDVVEITENSQELMRNSSGP	180	
Qy	181	CLSGSLVSLHCLACCKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLVT	240	
Db	181	CLSGSLVSLHCLACCKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSLDPHWLVT	240	
Qy	241	AAHCFRKHITDVENWVKVRAGSKLGSFPPLAVAKIIIEENPMYKXNDITAMKLOPPLTF	300	
Db	241	AAHCFRKHITDVENWVKVRAGSKLGSFPPLAVAKIIIEENPMYKXNDITAMKLOPPLTF	300	
Qy	301	SGTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLOASVQVVIDSTRCNADDAY	360	
Db	301	SGTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLOASVQVVIDSTRCNADDAY	360	


```
Qy 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Db 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Qy 421 SAYLNWIYNVWKAE 435
Db 421 SAYLNWIYNVWKAE 435

RESULT 2
US-10-473-127-1578
; Sequence 1578, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1578
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1578

Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.9e-219;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPDSQPLNSLDVVKPLRPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Db 1 MDPDSQPLNSLDVVKPLRPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Qy 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Db 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWILT 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWILT 240
Qy 241 AAHCFRKHDTVFNKVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Db 241 AAHCFRKHDTVFNKVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Qy 301 SGTVRPICLPFFDEELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Db 301 SGTVRPICLPFFDEELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Qy 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Db 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Qy 421 SAYLNWIYNVWKAE 435
Db 421 SAYLNWIYNVWKAE 435

RESULT 4
US-09-968-415-6
; Sequence 6, Application US/09968415
```

```
Db 421 SAYLNWIYNVWKAE 435

RESULT 3
US-10-473-127-1597
; Sequence 1597, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1597
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1597
```

```
Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 1.9e-219;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPDSQPLNSLDVVKPLRPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Db 1 MDPDSQPLNSLDVVKPLRPKPRIPMETFRKVGIPPIIALLSLASIIIVVLIKVILDKYF 60
Qy 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Db 61 LCGQPLHPIPRKQLCDGELDCPLGEDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
Qy 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Db 121 WFSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWILT 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSILDPHWILT 240
Qy 241 AAHCFRKHDTVFNKVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Db 241 AAHCFRKHDTVFNKVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLQFPLTF 300
Qy 301 SGTVRPICLPFFDEELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Db 301 SGTVRPICLPFFDEELTPATPLWIIWGFTKQNGKMSDILLQASVQVIDSTRCNADDA 360
Qy 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Db 361 QGEVTERKMCAGIPGEGVDTCCGDSGGPLMYQSDQHWVGVISWGYCGGSPSTPGVYTKV 420
Qy 421 SAYLNWIYNVWKAE 435
Db 421 SAYLNWIYNVWKAE 435
```

```
RESULT 4
US-09-968-415-6
; Sequence 6, Application US/09968415
```

```

; Publication No. US20020086334A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
;
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968,415
; FILING DATE: 26-Sep-2001
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/659,151
; FILING DATE: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
;
; INFORMATION FOR SEQ ID NO: 6:
;
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT13
; CLONE: 1337018
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-09-968-415-6
Query Match          99.8%; Score 2338; DB 3; Length 435;
Best Local Similarity 99.8%; Pred. No. 4.7e-219;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVVKPLKPRIPMETFRKVGPIPIIIALLSLASIIIVVVLKIVLDKYFF 60
DB 1 MDPDSQPLNSLDVVKPLKPRIPMETFRKVGPIPIIIALLSLASIIIVVVLKIVLDKYFF 60
QY 61 LCGQPLHFIPRKQLCDGELDCPLGBDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
DB 61 LCGQPLHFIPRKQLCDGELDCPLGBDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
QY 121 WFSACFDNFTALAEATACRQMGYSKPTTFAVEIGPDQDLVDVEITENSQELRMNNSGP 180
DB 121 WFSACFDNFTALAEATACRQMGYSKPTTFAVEIGPDQDLVDVEITENSQELRMNNSGP 180
QY 181 CLSGSLVSLHCLACGSKLTPRWGGEASVDSWPNQVSIQYDKHVCVCGSLDHPHWLT 240
DB 181 CLSGSLVSLHCLACGSKLTPRWGGEASVDSWPNQVSIQYDKHVCVCGSLDHPHWLT 240
QY 241 AAHCFRKHTDVENWVKRAGSDKLGSPSLAVAKIIIEBNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHTDVENWVKRAGSDKLGSPSLAVAKIIIEBNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPICLPFFDEBELTPATPLWIIIGWFTKQNGKMSDILLQASVQVLDSTRCNADDAY 360

; Publication No. US20030166246A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
;
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/180,719
; FILING DATE: 25-Jun-2002
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/008,271
; FILING DATE: 16-Jan-1998
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
;
; INFORMATION FOR SEQ ID NO: 6:
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT13
; CLONE: 1337018
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-10-180-719-6
Query Match          99.8%; Score 2338; DB 4; Length 435;
Best Local Similarity 99.8%; Pred. No. 4.7e-219;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVVKPLKPRIPMETFRKVGPIPIIIALLSLASIIIVVVLKIVLDKYFF 60
DB 1 MDPDSQPLNSLDVVKPLKPRIPMETFRKVGPIPIIIALLSLASIIIVVVLKIVLDKYFF 60
QY 61 LCGQPLHFIPRKQLCDGELDCPLGBDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
DB 61 LCGQPLHFIPRKQLCDGELDCPLGBDEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
QY 121 WFSACFDNFTALAEATACRQMGYSKPTTFAVEIGPDQDLVDVEITENSQELRMNNSGP 180
DB 121 WFSACFDNFTALAEATACRQMGYSKPTTFAVEIGPDQDLVDVEITENSQELRMNNSGP 180
QY 181 CLSGSLVSLHCLACGSKLTPRWGGEASVDSWPNQVSIQYDKHVCVCGSLDHPHWLT 240
DB 181 CLSGSLVSLHCLACGSKLTPRWGGEASVDSWPNQVSIQYDKHVCVCGSLDHPHWLT 240
QY 241 AAHCFRKHTDVENWVKRAGSDKLGSPSLAVAKIIIEBNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHTDVENWVKRAGSDKLGSPSLAVAKIIIEBNPMYPKNDIALMKLOPPLTF 300
QY 301 SGTVRPICLPFFDEBELTPATPLWIIIGWFTKQNGKMSDILLQASVQVLDSTRCNADDAY 360
```

Db 61 LCGQLHFIIPRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
Qy 121 WFSACDFNTEALAEATACQMGYSKPTFRFAVEIGPDQDLDDVVEITENSQELMRNSSGP 180
Db 121 WFSACDFNTEALAEATACQMGYSKPTFRFAVEIGPDQDLDDVVEITENSQELMRNSSGP 180
Qy 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
Db 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
Qy 241 AAHCFRKHDTVFNWKNVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
Db 241 AAHCFRKHDTVFNWKNVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
Qy 301 SGTVPICLPFDEELTPATLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
Db 301 SGTVPICLPFDEELTPATLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
Qy 361 QGEVTEKMWKAGIPGGVDTTCQDGGPLMYQSDQHWVGVISWGYCGGSPSTGVTYKV 420
Db 361 QGEVTEKMWKAGIPGGVDTTCQDGGPLMYQSDQHWVGVISWGYCGGSPSTGVTYKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435

RESULT 6
US-10-473-127-1596
; Sequence 1596, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026WO1
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1596
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1596

Query Match 99.8%; Score 2338; DB 5; Length 435;
Best Local Similarity 99.8%; Pred. No. 4.7e-219;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDPDSQPLNSLDVVKLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVLIKVLIDKYF 60
Db 1 MDPDSQPLNSLDVVKLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVLIKVLIDKYF 60
Qy 61 LCGQLHFIIPRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
Db 61 LCGQLHFIIPRKQKQDGLDCLPGBDEEHCVKSPFEGPAVAVRLSKDRSTLQVLDSATGN 120
Qy 121 WFSACDFNTEALAEATACQMGYSKPTFRFAVEIGPDQDLDDVVEITENSQELMRNSSGP 180
Db 121 WFSACDFNTEALAEATACQMGYSKPTFRFAVEIGPDQDLDDVVEITENSQELMRNSSGP 180

Qy 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
Db 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHHWLT 240
Qy 241 AAHCFRKHDTVFNWKNVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
Db 241 AAHCFRKHDTVFNWKNVRAGSKLGSFPFSLAVAKIIIEFNPMYPKNDIALMKLOFPLTF 300
Qy 301 SGTVPICLPFDEELTPATLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
Db 301 SGTVPICLPFDEELTPATLWIIIGWFTKQNGKMSDILLOASVQVIDSTRCNADAY 360
Qy 361 QGEVTEKMWKAGIPGGVDTTCQDGGPLMYQSDQHWVGVISWGYCGGSPSTGVTYKV 420
Db 361 QGEVTEKMWKAGIPGGVDTTCQDGGPLMYQSDQHWVGVISWGYCGGSPSTGVTYKV 420
Qy 421 SAYLNWIYNWKAEL 435
Db 421 SAYLNWIYNWKAEL 435
RESULT 7
US-11-045-577-6
; Sequence 6, Application US/11045577
; Publication No. US20050181404A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/045,577
; FILING DATE: 27-Jan-2005
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/659,151
; FILING DATE: 11-Sep-2000
; APPLICATION NUMBER: 09/008,271
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PF-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 6 :
US-11-045-577-6

```
Query Match          99.8%; Score 2338; DB 6; Length 435;
Best Local Similarity 99.8%; Pred. No. 4,7e-219;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVPLRKPRIPIIPIIALLSLASIIIVVLKIVLDKYIF 60
DB 1 MDPDSQPLNSLDVPLRKPRIPIIPIIALLSLASIIIVVLKIVLDKYIF 60

QY 61 LCGQPLHFIIPKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLDATGN 120
DB 61 LCGQPLHFIIPKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLDATGN 120

QY 121 WFSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180
DB 121 WFSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 180

QY 181 CLSGSLVSLHCLACGSKSLKTRPVVGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLTA 240
DB 181 CLSGSLVSLHCLACGSKSLKTRPVVGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLTA 240

QY 241 AAHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300
DB 241 AAHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 300

QY 301 SGTVPICLPFPDEBLTATPLWIIGWFTKONGKMSDILLOASVQVLDSTRCNADAY 360
DB 301 SGTVPICLPFPDEBLTATPLWIIGWFTKONGKMSDILLOASVQVLDSTRCNADAY 360

QY 361 QGEVTEKMCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKV 420
DB 361 QGEVTEKMCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKV 420

QY 421 SAYLWNIYNVWKAEL 435
DB 421 SAYLWNIYNVWKAEL 435

RESULT 8
US-09-851-588-8
; Sequence 8, Application US/09851588
; Patent No. US20020042067A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR FILING DATE: 2000-08-17
; PRIOR FILING DATE: 2000-08-17
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-8

Query Match          99.8%; Score 2337; DB 3; Length 437;
Best Local Similarity 100.0%; Pred. No. 5,9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPLRKPRIPIIPIIALLSLASIIIVVLKIVLDKYIF 61
DB 4 DPDSQPLNSLDVPLRKPRIPIIPIIALLSLASIIIVVLKIVLDKYIF 63

QY 62 CGQPLHFIIPKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLDATGN 121
```

```
DB 64 CGQPLHFIIPKQLCDGELDCPLGDEBHCVKSPFEGPAVAVRLSKDRSTLQVLDATGN 123
QY 122 FSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 181
DB 124 FSACDFNFTALAEATACRQMGYSKPTFRAVEIGPDQDLDVVEITENSQELMRNSSGP 183
QY 182 LSGSLVSLHCLACGSKSLKTRPVVGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLTA 241
DB 184 LSGSLVSLHCLACGSKSLKTRPVVGEEASVDSWPQVSIQYDKQHVCGGSLDHPHWLTA 243
QY 242 AHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 301
DB 244 AHCFRKHDTVFNWVKRAGSKLGFSPSLAVAKIIIEFNPMYPKNDIALMKLOPPLTF 303
QY 302 GTVPICLPFPDEBLTATPLWIIGWFTKONGKMSDILLOASVQVLDSTRCNADAY 361
DB 304 GTVPICLPFPDEBLTATPLWIIGWFTKONGKMSDILLOASVQVLDSTRCNADAY 363
QY 362 GEVTEKMCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKV 421
DB 364 GEVTEKMCAGIPGGVDTCQDGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKV 423
QY 422 AYLWNIYNVWKAEL 435
DB 424 AYLWNIYNVWKAEL 437

RESULT 9
US-10-295-027-779
; Sequence 779, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIOR FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-09-15
; PRIOR FILING DATE: 2000-09-15
; PRIOR FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2001-11-13
; PRIOR FILING DATE: 2001-11-15
; PRIOR FILING DATE: 2001-11-15
; PRIOR FILING DATE: 2001-11-21
; PRIOR FILING DATE: 2001-11-21
; PRIOR FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2001-11-29
; PRIOR FILING DATE: 2001-12-14
; PRIOR FILING DATE: 2001-12-14
; PRIOR FILING DATE: 2002-01-08
; PRIOR FILING DATE: 2002-01-08
; PRIOR FILING DATE: 2002-01-10
; PRIOR FILING DATE: 2002-01-10
; PRIOR FILING DATE: 2002-02-08
; PRIOR FILING DATE: 2002-02-08
; PRIOR FILING DATE: 2002-02-13
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 779
; LENGTH: 437
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-295-027-779

Query Match
Best Local Similarity 99.8%; Score 2337; DB 4; Length 437;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVLIDKYFL 61
Db 4 DPDSQPLNSLDVPRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVLIDKYFL 63
QY 62 CGQPLHFIPRKQDCGELDCPLGDEEHCVKSPPEGPAVAVRLSKDRSTLQVLDSATGNW 121
Db 64 CGQPLHFIPRKQDCGELDCPLGDEEHCVKSPPEGPAVAVRLSKDRSTLQVLDSATGNW 123
QY 122 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSLELRNNSGPG 181
Db 124 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSLELRNNSGPG 183
QY 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 243
QY 242 AHCFRKHDTVFNWVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 301
Db 244 AHCFRKHDTVFNWVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 303
QY 302 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAYQ 361
Db 304 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAYQ 363
QY 362 GEVTEKMCAGIPEGGVDTCCGDSGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKVS 421
Db 364 GEVTEKMCAGIPEGGVDTCCGDSGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKVS 423
QY 422 AYLNWIYNNWKAEL 435
Db 424 AYLNWIYNNWKAEL 437

RESULT 10
US-10-295-027-791
; Sequence 791, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
```

```
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 791
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-791

Query Match
Best Local Similarity 99.8%; Score 2337; DB 4; Length 437;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVPRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVLIDKYFL 61
Db 4 DPDSQPLNSLDVPRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVLIDKYFL 63
QY 62 CGQPLHFIPRKQDCGELDCPLGDEEHCVKSPPEGPAVAVRLSKDRSTLQVLDSATGNW 121
Db 64 CGQPLHFIPRKQDCGELDCPLGDEEHCVKSPPEGPAVAVRLSKDRSTLQVLDSATGNW 123
QY 122 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSLELRNNSGPG 181
Db 124 FSACFDNFTALAEACROMGYSSKPTFRAVEIGPDQDLDVVEITENSLELRNNSGPG 183
QY 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 241
Db 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSILDPHWLTA 243
QY 242 AHCFRKHDTVFNWVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 301
Db 244 AHCFRKHDTVFNWVKVRAGSKLGSFSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 303
QY 302 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAYQ 361
Db 304 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAYQ 363
QY 362 GEVTEKMCAGIPEGGVDTCCGDSGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKVS 421
Db 364 GEVTEKMCAGIPEGGVDTCCGDSGGPLMYQSDQHVHVGIVSWGCGGPGSTPGVYTKVS 423
QY 422 AYLNWIYNNWKAEL 435
Db 424 AYLNWIYNNWKAEL 437

RESULT 11
US-10-295-027-831
; Sequence 831, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
```

APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: 2002-11-13
PRIORITY APPLICATION NUMBER: US 09/663,733
PRIORITY FILING DATE: 2000-09-15
PRIORITY APPLICATION NUMBER: US 60/350,666
PRIORITY FILING DATE: 2001-11-13
PRIORITY APPLICATION NUMBER: US 60/335,394
PRIORITY FILING DATE: 2001-11-15
PRIORITY APPLICATION NUMBER: US 60/332,464
PRIORITY FILING DATE: 2001-11-21
PRIORITY APPLICATION NUMBER: US 60/334,393
PRIORITY FILING DATE: 2001-11-29
PRIORITY APPLICATION NUMBER: US 60/340,376
PRIORITY FILING DATE: 2001-12-14
PRIORITY APPLICATION NUMBER: US 60/347,211
PRIORITY FILING DATE: 2002-01-08
PRIORITY APPLICATION NUMBER: US 60/347,349
PRIORITY FILING DATE: 2002-01-10
PRIORITY APPLICATION NUMBER: US 60/355,250
PRIORITY FILING DATE: 2002-02-13
PRIORITY APPLICATION NUMBER: US 60/356,714
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 831
LENGTH: 437
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-831

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVKEPRIPMETFRKVGIPIIIIALLSLASIIIVVILKILDKYPL 61
DB 4 DPDSQPLNSLDVKEPRIPMETFRKVGIPIIIIALLSLASIIIVVILKILDKYPL 63
QY 62 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 121
DB 64 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 123
QY 122 FSACFDNTEALAEACRMQYSSKPTFRAVIGDPDQDLVVETIENSQELMRNSSGPC 181
DB 124 FSACFDNTEALAEACRMQYSSKPTFRAVIGDPDQDLVVETIENSQELMRNSSGPC 183
QY 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWLTA 241
DB 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWLTA 243
QY 242 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTFPS 301
DB 244 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTFPS 303
QY 302 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
DB 304 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
QY 362 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVGVGIVSWGCGGSPSTPGVYTKVS 421
DB 364 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVGVGIVSWGCGGSPSTPGVYTKVS 423
QY 422 AYLNMIYNWKAEL 435
DB 424 AYLNMIYNWKAEL 437

RESULT 12

US-10-295-027-1196
Sequence 1196, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afari, Daniel
APPLICANT: Aziz, Natacha

APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT FILING DATE: 2002-11-13
PRIORITY APPLICATION NUMBER: US 09/663,733
PRIORITY FILING DATE: 2000-09-15
PRIORITY APPLICATION NUMBER: US 60/350,666
PRIORITY FILING DATE: 2001-11-13
PRIORITY APPLICATION NUMBER: US 60/335,394
PRIORITY FILING DATE: 2001-11-15
PRIORITY APPLICATION NUMBER: US 60/332,464
PRIORITY FILING DATE: 2001-11-21
PRIORITY APPLICATION NUMBER: US 60/334,393
PRIORITY FILING DATE: 2001-11-29
PRIORITY APPLICATION NUMBER: US 60/340,376
PRIORITY FILING DATE: 2001-12-14
PRIORITY APPLICATION NUMBER: US 60/347,211
PRIORITY FILING DATE: 2002-01-08
PRIORITY APPLICATION NUMBER: US 60/347,349
PRIORITY FILING DATE: 2002-01-10
PRIORITY APPLICATION NUMBER: US 60/355,250
PRIORITY FILING DATE: 2002-02-08
PRIORITY APPLICATION NUMBER: US 60/356,714
PRIORITY FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1196
LENGTH: 437
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1196

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVKEPRIPMETFRKVGIPIIIIALLSLASIIIVVILKILDKYPL 61
DB 4 DPDSQPLNSLDVKEPRIPMETFRKVGIPIIIIALLSLASIIIVVILKILDKYPL 63
QY 62 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 121
DB 64 CGQPLHFTPRKQDCGELDCPLGEDEEHCHVKSPFEGPAVAVRLSKDRSTLQVLDATGNW 123
QY 122 FSACFDNTEALAEACRMQYSSKPTFRAVIGDPDQDLVVETIENSQELMRNSSGPC 181
DB 124 FSACFDNTEALAEACRMQYSSKPTFRAVIGDPDQDLVVETIENSQELMRNSSGPC 183
QY 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWLTA 241
DB 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIILDPHWLTA 243
QY 242 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTFPS 301
DB 244 AHCPRKHTDVFNNKVRAGSKLGSFPLSAVAKIIIEFNPMYPKNDIALMKLQPLTFPS 303
QY 302 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
DB 304 GTVRPCLPFFDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
QY 362 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVGVGIVSWGCGGSPSTPGVYTKVS 421
DB 364 GEVTEKMMKAGIPEGGVDTCCQDSGGPLMYQSDQHVGVGIVSWGCGGSPSTPGVYTKVS 423

Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437

RESULT 13

US-10-173-999-89
; Sequence 89, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 89
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-89

Query Match 99.8%; Score 2337; DB 4; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DPDSQPLNSLDVVKLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDKYFL 61
Db 4 DPDSQPLNSLDVVKLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDKYFL 63
Qy 62 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSFPEGPVAVRLSKDRSTLQVLD SATGNW 121
Db 64 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSFPEGPVAVRLSKDRSTLQVLD SATGNW 123
Qy 122 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPC 181
Db 124 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPC 183
Qy 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 241
Db 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 243
Qy 242 AHCFRKHDTDFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 301
Db 244 AHCFRKHDTDFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 244 AHCFRKHDTDFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 304 GTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
Qy 362 GEVTEKMWKAGIPGGVDTCGDSGGPLMYOSDQHVHVGIVSWGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMWKAGIPGGVDTCGDSGGPLMYOSDQHVHVGIVSWGCGGSPSTPGVYTKVS 423
Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437

RESULT 15

US-10-473-127-1586
; Sequence 1586, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.

RESULT 14
US-10-473-127-1581
; Sequence 1581, Application US/10473127
; Publication No. US20040236091A1
; GENERAL INFORMATION:
; APPLICANT: Zycos Inc.
; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1581
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1581

Query Match 99.8%; Score 2337; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DPDSQPLNSLDVVKLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDKYFL 61
Db 4 DPDSQPLNSLDVVKLRKPRIPMETFRKVGIPPIIIALLSLASIIIVVVLKVIDKYFL 63
Qy 62 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSFPEGPVAVRLSKDRSTLQVLD SATGNW 121
Db 64 CGQPLHFI PRKQLCDGELDCPLGEDEEHCVKSFPEGPVAVRLSKDRSTLQVLD SATGNW 123
Qy 122 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPC 181
Db 124 FSACPDNFTALAEATACROMGYSSKPTFRAVEIGPDQDLVVEITENSQELRMNSSGPC 183
Qy 182 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 241
Db 184 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 243
Qy 242 AHCFRKHDTDFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 301
Db 244 AHCFRKHDTDFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTFS 303
Qy 302 GTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 361
Db 304 GTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADDAVQ 363
Qy 362 GEVTEKMWKAGIPGGVDTCGDSGGPLMYOSDQHVHVGIVSWGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMWKAGIPGGVDTCGDSGGPLMYOSDQHVHVGIVSWGCGGSPSTPGVYTKVS 423
Qy 422 AYLNWIYNVWKAEL 435
Db 424 AYLNWIYNVWKAEL 437


```

; TITLE OF INVENTION: TRANSLATIONAL PROFILING
; FILE REFERENCE: 08191-026W01
; CURRENT APPLICATION NUMBER: US/10/473,127
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 60/279,495
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: 60/292,544
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 60/310,801
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: 60/326,370
; PRIOR FILING DATE: 2001-10-01
; PRIOR APPLICATION NUMBER: 60/336,780
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/358,985
; PRIOR FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 2041
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1586
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-473-127-1586

```

```

Query Match      99.8%; Score 2337; DB 5; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.9e-219;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DPDSQPLNSLDVKPLRPKIPMETFRKVGPIPIIIALLSLASIIIVVLIKVILDKYYFL 61
Db 4 DPDSQPLNSLDVKPLRPKIPMETFRKVGPIPIIIALLSLASIIIVVLIKVILDKYYFL 63

Qy 62 CGQPLHFIIPKQLCDGELDCPLGEDEHCVKSFPEGPAAVAVRLSKDRSTLQVLD SATGNW 121
Db 64 CGQPLHFIIPKQLCDGELDCPLGEDEHCVKSFPEGPAAVAVRLSKDRSTLQVLD SATGNW 123

Qy 122 FSACFDNTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGPC 181
Db 124 FSACFDNTEALAEATACRQMGYSKPTFRAVEIGPDQDLVDVEITENSQELMRNSSGPC 183

Qy 182 LSGSLVSLHCLACGSKLTPTRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 241
Db 184 LSGSLVSLHCLACGSKLTPTRVVGEEASVDSWPQVSIQYDKQHVCGGSIIDPHWVLT 243

Qy 242 AHCPRKHTDVFNWKVRAGSKLGSPPSLAVAKIIIEFNPMYPKDNIDIALMKLQFPLTFS 301
Db 244 AHCPRKHTDVFNWKVRAGSKLGSPPSLAVAKIIIEFNPMYPKDNIDIALMKLQFPLTFS 303

Qy 302 GTVRPCLPFDEELTPATPLMIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAYQ 361
Db 304 GTVRPCLPFDEELTPATPLMIIGWFTKQNGKMSDILLOASVQVIDSTRCNADDAYQ 363

Qy 362 GEVTEKMMKAGIPEGGVDTCCGDSGGPLMYQSDQHHVVGIVSWGYCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMMKAGIPEGGVDTCCGDSGGPLMYQSDQHHVVGIVSWGYCGGSPSTPGVYTKVS 423

Qy 422 AYLNIWYINWKAEL 435
Db 424 AYLNIWYINWKAEL 437

```

Search completed: December 5, 2005, 14:07:34
Job time : 168 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 20:06:35 ; Search time 236 Seconds
(without alignments)
3276.438 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVKPLRKP.....VYTKSVLNIWVNWKAEL 435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10803530/runat_05122005_083216_2147/app.query.fasta_1.583
-DB=Issued Patents NA -OPWT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10803530@cgn_1_1_290 @runat_05122005_083216_2147 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/1 COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
7: /cgn2_6/ptodata/1/ina/pp COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2338	99.8	2038	3	US-09-008-271A-18
2	2338	99.8	2038	3	US-09-968-415-18
3	2337	99.8	2081	3	US-09-851-588-7
4	2335	99.7	2121	3	US-09-607-745-1
5	2324	99.2	2079	3	US-09-656-002-1
6	2324	99.2	2079	3	US-09-851-588-5
7	2297.5	98.1	2063	3	US-10-012-231A-274
8	2297.5	98.1	2063	3	US-10-015-389A-274
9	2297.5	98.1	2063	3	US-10-006-768A-274

10	2297.5	98.1	2063	3	US-10-015-671A-274
11	2297.5	98.1	2063	3	US-10-015-333A-274
12	2297.5	98.1	2063	3	US-10-011-833A-274
13	2297.5	98.1	2063	3	US-10-006-041A-274
14	2297.5	98.1	2063	3	US-10-012-064A-274
15	1259	53.8	1189	3	US-09-607-745-8
16	698.5	29.8	1434	3	US-10-177-661-1
17	697.5	29.8	1748	3	US-09-879-792-11
18	683	29.2	1341	3	US-10-177-661-3
19	677.5	28.9	2479	3	US-09-342-749-29
20	677.5	28.9	2479	3	US-09-691-840-29
21	677.5	28.9	2479	3	US-09-685-166A-894
22	677.5	28.9	2479	3	US-09-679-426-894
23	677.5	28.9	2479	3	US-09-759-143-894
24	677.5	28.9	2479	3	US-10-012-896-894
25	676.5	28.9	1479	3	US-09-342-749-1
26	676.5	28.9	1479	3	US-09-691-840-1
27	676.5	28.9	2448	3	US-09-949-016-5203
28	675.5	28.8	1476	3	US-09-759-143-931
29	675.5	28.8	1476	3	US-10-012-896-931
30	675.5	28.8	1479	3	US-09-759-143-930
31	675.5	28.8	1479	3	US-10-012-896-930
32	675.5	28.8	3245	3	US-09-759-143-929
33	675.5	28.8	3245	3	US-10-012-896-929
34	663.5	28.3	3443	3	US-09-807-201-8
35	660.5	28.2	2440	3	US-09-949-016-5210
36	660.5	28.2	2440	3	US-09-949-016-5211
37	660.5	28.2	2440	3	US-09-949-016-5212
38	658	28.1	2412	3	US-09-999-833A-68
39	658	28.1	2412	3	US-10-020-445A-68
40	650.5	27.8	2413	3	US-09-518-046-1
41	650.5	27.8	2413	3	US-09-650-371-1
42	628.5	26.8	2070	3	US-09-856-371A-7
43	627	26.8	2544	3	US-09-518-046-3
44	627	26.8	2544	3	US-09-650-371-3
45	622.5	26.6	1230	3	US-09-879-792-35

ALIGNMENTS

RESULT 1

US-09-008-271A-18
; Sequence 18, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/09/008,271A
; FILING DATE: 16-Jan-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:

```

; NAME: Mohan-Peterson, Sheela
; REGISTRATION NUMBER: 41,201
; REFERENCE/DOCKET NUMBER: PP-0458 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2038 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: COLNOT13
; CLONE: 1337018
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-008-271A-18

Alignment Scores:
Pred. No.:      8,646-237      Length:      2038
Score:          2338.00      Matches:      434
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.77%      Mismatches: 0
Query Match:      99.83%      Indels:      0
Dbs:              3          Gaps:      0

US-10-803-530-2 (1-435) x US-09-008-271A-18 (1-2038)

Qy      1 MetAspProAspSerAspGluProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
Db      200 ATGGATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTCGCAACCC 259

Qy      21 ArgLeuProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      260 CGTATCCCATGGAGACCTCTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGC 319

Qy      41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrPhe 60
Db      320 CTGGCGAGTATCATATTGTGTGCTCTCATCAAGGTGATTCTGGATAAATACTACTTC 379

Qy      61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
Db      380 CTCTGGGGAGCCCTCTCCATTTATCCGAGGAGCAGCTGTGTGACGGAGAGCTGGAC 439

Qy      81 CysProLeuGlyGluAspGluLysCysValLysSerPheProGluGlyProAlaVal 100
Db      440 TGTCCCTTGGGGAGGACGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCTGCAAGTG 499

Qy      101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      500 GCAGTCCGCCCTCTCCAAAGGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGGAAC 559

Qy      121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysAtgGln 140
Db      560 TGGTTCTCTGCTGTTTCGCAACTTCACAGAGCTCTCGCTGAGACAGGCTGTAGGCAG 619

Qy      141 MetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLeu 160
Db      620 ATGGGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGGATCTG 679

Qy      161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      680 GATGTTGTTGNAATCAGAAACAGCCAGAGCTTCGCATGCGGAACTCAAGTGGGGCCC 739

Qy      181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      740 TGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTCCCTGTGGGAGAGCCTGAAGACC 799

Qy      201 ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      800 CCCCCTGTGGTGGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTCAGCATC 859

Qy      221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240

```

```

Db      860 CAGTACGACAAACAGCAGCTCTGTGGAGGAGCATCTCGACCCCACTGGTCTCTCACG 919
Qy      241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      920 GCAGCCCACTGCTTCAGAAACATACCGATGTGTTCACTGGAGAGTGGCGGAGGCTCA 979
Qy      261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      980 GACAAACTGGGCAGCTTCCCATCTCCCTGGCTGTGCCCAAGATCATCATTTGAATTCAAC 1039
Qy      281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db      1040 CCCATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCACCACTCCTTC 1099
Qy      301 SerGlyThrValArgProIleCysLeuProPheAspGluLeuThrProAlaThr 320
Db      1100 TCAGGCACAGTCAGGCCCATCTCTCTGCTCTTCTTGTATGAGGAGCTCACTCCAGCCACC 1159
Qy      321 ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db      1160 CCACCTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGGAGGAAGATGTCTGACATA 1219
Qy      341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db      1220 CTGCTGCAGGCGTCAGTCCAGGTCATTTGACGACACAGCGTGCATGCAGACATGCGTAC 1279
Qy      361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db      1280 CAGGGGAAATTCACCCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACC 1339
Qy      381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db      1340 TGCAGGGTGACAGTGGTGGGCCCTGTGTACCAATCTGACAGTGGCATGTGGTGGGC 1399
Qy      401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db      1400 ATCGTTAGCTGGGCTATGCTCGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTC 1459
Qy      421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1460 TCAGCCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1504

RESULT 2
US-09-968-415-18
; Sequence 18, Application US/09968415
; Patent No. 6855811
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968,415
; FILING DATE: 26-Sep-2001
; PRIOR APPLICATION DATA:

```


US-09-851-588-7

Alignment Scores:

Pred. No.: 1,14e-236 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-851-588-7 (1-2081)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 224 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAAAACCCGT 283
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuSerLeu 41
DB 284 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTG 343
QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
DB 344 GCGAGTATCATATTGTGTCTCATCAAGGTGATCTCGATAAATACTACTTCTC 403
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
DB 404 TCGGGGAGCCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGAGAGCTGGACTGT 463
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
DB 464 CCTTTGGGGAGGACGAGGACACTGTGTCAAGAGCTTCCCGAAGGCCCTGCGAGTGCA 523
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
DB 524 GTCCGCCCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGGAAGTGG 583
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 584 TTCTCTGCTGTTCGACAACTTCACAGAACTCTCGCTGAGACAGCTGTGAGCAGATG 643
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 644 GGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGAGATTCGCCAGACCCAGGATCTGGAT 703
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 704 GTTGTGTAATACAGAAAAACAGCAGGAGCTTCGCATCGGAACTCAAGTGGGCCCTGT 763
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 764 CTCTCAGGCTCCCTGGTCTCCCTGCATCTGTTCCTGTGGGAGAGCCCTGAAGACCCCC 823
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 824 CGTGTGTGGTGGGAGGAGCCCTGTGTGATTTCTTGGCTTGGCAGGTTCAGCATCCAG 883
QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 884 TACGACAAACAGCAGCTGTGAGGAGGAGCATCTGGACCCCACTGGGTCTCTCAGGCA 943
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
DB 944 GCCCACTGCTTCAGAAACATACCGATGTGTTCNACTGGAAAGGTGGGGAGGCTCAGAC 1003
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
DB 1004 AAATGGGAGCTTCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCC 1063
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1064 ATGTACCCCAAGACAAATGATCGCCCTCATGAAGCTGCAGTTCACACTCACTTTCTCA 1123
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321

DB 1124 GGCACAGTCAGGCCCATCTGTCTGCCCTTTTGTATGAGGAGCTCACTCCAGCCACCCCA 1183
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
DB 1184 CTCTGGATCATTTGATGGGCTTTTACGAAGCAGAAATGGAGGAAGATGTCTGACATACTG 1243
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1244 CTGAGGGCTCAGTCCAGGTCATTTGACAGCACACGGTGCATTCAGACAGCATGCTGAC 1303
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1304 GGGGAAGTCACCCGAGAGATGATGTGTGCGAGCATCCCGAAGGGGTGTGGACACTGC 1363
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
DB 1364 CAGGGTGACAGTGTGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGTGGGCATC 1423
QY 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
DB 1424 GTTAGCTGGGCTATGGCTCGGGGGCCCGAGCACCCAGAGTATACACCAAGGTCTCA 1483
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1484 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1525

RESULT 4

US-09-607-745-1
; Sequence 1, Application US/09607745
; Patent No. 6750034
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-607-745-1

Alignment Scores:

Pred. No.: 1,91e-236 Length: 2121
Score: 2335.00 Matches: 434
Percent Similarity: 99.77% Conservativeness: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 99.70% Indels: 0
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-607-745-1 (1-2121)

QY 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysPro 20
DB 277 ATGGATCTTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCAAAACCC 336
QY 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
DB 337 CGTATCCCATCGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCATCTAGC 396
QY 41 LeuAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTrpPhe 60
DB 397 CTGGCGAGTATCATATTGTGTGTCTCATCAAGGTGATTCGATAAAATACTACTTC 456
QY 61 LeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAsp 80
DB 457 CTCTCGGGCAGCCTCTCCACTTTCATCCGAGGAAGAGCTGTGTGACGGAGCTGGAC 516

```
QY      81  CysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      517  TGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTGCAGTG 576

QY      101  AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      577  GCAGTCGGCTCTCCAAAGACCGATCCACACTGCAGTGTGTGACTCGGCCACAGGGAAC 636

QY      121  TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db      637  TGGTTCTCTGCTGTGTTCACAACATTCACAGAAGCTCTCGCTGAGACAGCTGTAGGCAG 696

QY      141  MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
Db      697  ATGGGCTACAGACGACAAACCCATTTTCAGAGCTGTGGAGATTGGCCACAGCAGGATCTG 756

QY      161  AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      757  GATGTGTTGTAATCACAGAAACAGCCAGGAGCTTCGCATCGGGAACCTCAAGTGGGCC 816

QY      181  CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      817  TGTCTCTCAGGCTCCCTGCTCCCTGCACCTGTCTTGCCTGTGGGAAGAGCCTGAAGACC 876

QY      201  ProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      877  CCCCCTGTGTGGTGGGAGGAGGCTCTGTGGATTCTTGGCTTGGCAGGTTCAGATC 936

QY      221  GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db      937  CAGTACGACAAACAGCACCTCTGTGGAGGAGCATCTCGACCCCTCAGTGGGTCTCAGC 996

QY      241  AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      997  GCAGCCCACTGTCTCAGGAACATACCATGTGTTCAACTGGAAGTGGCGGCGAGCTCA 1056

QY      261  AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
Db      1057  GACAACTGGGAGCTTCCATCCCTGGCTGTGGCCCAAGATCATCATCAATGAATCAAC 1116

QY      281  ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db      1117  CCCATGTACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTCACATCACTTTC 1176

QY      301  SerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThr 320
Db      1177  TCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACC 1236

QY      321  ProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db      1237  CCACCTCTGGATCATTGGATGGGGCTTTACGAAAGCAGAAATGGAGGGAAGATGTCTGACATA 1296

QY      341  LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db      1297  CTGCTGCAGCGTCACTCAGGTCAATGACGACACGGTGCATGCAATGACAGCAGATGGGTAC 1356

QY      361  GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db      1357  CTGGGGGAAGTCACCGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACC 1416

QY      381  CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
Db      1417  TGCACGGGTGACAGTGTGGGCCCTTGATGTACCAATCTGACCCAGTGGCATGTGTGGGGC 1476

QY      401  IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
Db      1477  ATCGTTAGCTGGGGCTATGGCTCGGGGGGCCCGGAGCACCCCGAGGGGTATACCAAGGTC 1536

QY      421  SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1537  TCAGCCATCTCACTGGATCTACATGTCTGGAAGGCTGAGCTG 1581
```

RESULT 5

```
US-09-656-002-1
; Sequence 1, Application US/09656002
; Patent No. 6455668
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt
; APPLICANT: Wilson, Keith
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-69108/DJB/JUD/AMS
; CURRENT APPLICATION NUMBER: US/09/656,002
; CURRENT FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: US 09/525,993
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: US 09/493,444
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: PCT/US 00/07044
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2079
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-656-002-1
```

```
Alignment Scores:
Pred. No.:      2,68e-235      Length:      2079
Score:          2324.00        Matches:    434
Percent Similarity: 99.77%    Conservative: 0
Best Local Similarity: 99.77% Mismatches:    0
Query Match:      99.23%      Indels:      1
DB:               3           Gaps:        0
```

US-10-803-530-2 (1-435) x US-09-656-002-1 (1-2079)

```
QY      2  AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db      217  GATCTCGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCGCAACCCCTG 276

QY      22  IleProMetGluThrPheArgLys-ValGlyIleProIleIleIleAlaLeuLeuSerLe 41
Db      277  ATCCCCATGGAGACCTTCAGAAAGTGTGGGGATCCCATCATCATAGCACTACTGAGCCT 336

QY      41  uAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLe 61
Db      337  GCGAGATATCATTTGTGTGTCTCATCAAGGTGATTCGTGGATAAATACTACTTCTCT 396

QY      61  uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db      397  CTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTG 456

QY      81  sProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db      457  TCCTTTGGGGAGGAGAGGACATGTGTCAAGAGCTTCCCGAAGGCCCTGCACTGGC 516

QY      101  aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db      517  AGTCCGCTCTCCAAGGACCGATCCACACTGCAGGTGTGTGGACTCGGCCACAGGGAACCTG 576

QY      121  pPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMe 141
Db      577  GTTCTCTGCTCTTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGAT 636

QY      141  tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db      637  GGGCTACAGCAGCAAAACCCACTTTTCAGAGCTGTGGAGATTGGCCACAGACAGATCTGGA 696

QY      161  pValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCy 181
Db      697  TGTGTGTGAATCACAGAAACAGCCAGGAGCTTCGATCGCGGAACCTCAAGTGGGGCCCTG 756

QY      181  sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201
```


Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTTGCTGTGGAGAGAGCTGAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleG1 221
Db 817 CCGTGTGGTGGTGGGAGAGAGCCCTCTGTGGATTCTTGGCCCTGGCAGTCAAGTCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACACGACGTCGTGTGGAGGAGCATCTTGGACCCCTGGCTGGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAs 261
Db 937 AGCCCACTGCTTCAGGAAACATACCGATGTTTCAACTGGAAGGTGGCGCAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCATCCTGGCTGTGGCCAAAGATCATCATGTGAATTCAACCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTCTC 1116
QY 301 rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 1117 AGCCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGGAGCTCACTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 1177 ACTCTGGATCATTTGATGGGCTTTACGACGACAGATGGAGGAAGATGTCGTGACATCT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrG1 361
Db 1237 GCTGCAGGCGTCAGTCCAGGTCATTCACACACACCGTGCAATGCAGACGATGGTACCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAspThrCy 381
Db 1297 GGGGGAAGTCCAGAGAGATGATGTGCAGGATCCCGAAGGGGTGTGGACACCTG 1356
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 1357 CCAGGTCACAGTGTGTGGCCCTGATGATACCAATCTGACCATGGCATGTGTGGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSe 421
Db 1417 CTTAGCTGGGCTATGGCTCGGGGGCCGAGCACCCAGAGTATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCCTATCTCACTGGATCTACATGTCGTGAAGGCTGAGCTG 1519

RESULT 6

US-09-851-588-5
; Sequence 5, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DJB/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851.588
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2079
; TYPE: DNA

! ORGANISM: Homo sapiens
US-09-851-588-5

Alignment Scores:
Pred. No.: 2,686-235 Length: 2079
Score: 2324.00 Matches: 434
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.23% Indels: 1
DB: 3 Gaps: 0

US-10-803-530-2 (1-435) x US-09-851-588-5 (1-2079)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 217 GATCCTGACAGTGTATCAACCTCTGAACAGGCTCGATGTCAACCCCTGCGCAACCCCGT 276
QY 22 IleProMetGluThrPheArgLys-ValGlyIleProIleIleIleAlaLeuSerLe 41
Db 277 ATCCCATGGAGACCTTCAGAAAGTGGGGATCCCCATCATATAGCACTACTGAGCCT 336
QY 41 uAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLe 61
Db 337 GCGCAGTATCATATTGTGTCTCTCATCAAGGTGATCTCGATAAATACTACTTCT 396
QY 61 uCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCy 81
Db 397 CTGGGGAGCCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTACGGAGAGCTGGACTG 456
QY 81 sProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAl 101
Db 457 TCCCTTGGGGAGGAGCAGGAGACCTGTGTCAAGAGCTTCCCGAAGGGCTCGAGTGGC 516
QY 101 aValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTr 121
Db 517 AGTCCGCTCTCCAAAGACCGATCCACACTGCAGGTGTGGACTCGGCCACAGGAACTG 576
QY 121 pPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMe 141
Db 577 GTTCTCTGCTGTTCGACAACCTTCAGAAAGCTCTCGCTGAGACAGCTGTAGGCAGAT 636
QY 141 tGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAs 161
Db 637 GGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGCCCCAGACACAGGATCTGA 696
QY 161 pValValGluIleThrGluAsnSerGlnLeuArgMetArgAsnSerSerGlyProCy 181
Db 697 TGTGTGTGAATTCACAGAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTG 756
QY 181 sLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPr 201
Db 757 TCTCTCAGGCTCCCTGGTCTCCCTGCACCTGCTTGGCTGTGGAAAGAGCTGAAGACCCC 816
QY 201 oArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleG1 221
Db 817 CCGTGTGGTGGTGGGAGGAGGCTCTGTGATTCTTGGCCCTTGGCAGTCAAGTCCA 876
QY 221 nTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAl 241
Db 877 GTACGACAAACAGCAGCTGTGTGGAGGAGCATCTCGGACCCCTGCGGTCTCACGGC 936
QY 241 aAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAs 261
Db 937 AGCCCACTGCTTCAGGAAACATACCGATGTTTCAACTGGAAGGTGGCGCAGGCTCAGA 996
QY 261 pLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPr 281
Db 997 CAAACTGGGAGCTTCCCATCCTGGCTGTGGCCAAAGATCATCATGTGAATTCAACCC 1056
QY 281 oMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
Db 1057 CATGTACCCCAAGACAATGACATGCGCCCTCATGAGGCTGAGTTCCTCCACTCCTTCTC 1116

```
QY 301 rGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPr 321
Db 1117 AGGCACAGTCAGGCCCATCTGTCGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCC 1176
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 1177 ACTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGGAAGATGCTGTACATACT 1236
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaThrG1 361
Db 1237 GCTGAGGCGTCAGTCCAGGTCATTGACAGCACACGGTCAATGACAGCGATGCGTACCA 1296
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 1297 GGGGGAAGTCACCGAGAGATGATGTGTCAGGCAATCCCGAAGGGGTGTGCACACCTG 1356
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 1357 CCAGGGTGACAGTGTGGGCCCTCGATGTACCAATCTGACCAGTGGCATGTGTGGGCAT 1416
QY 401 eValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSe 421
Db 1417 CGTTAGCTGGGGCTATGGCTGGGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTC 1476
QY 421 rAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1477 AGCTATCTCAACTGATCTACAACTGTCTGGAAGGCTGAGCTG 1519
```

RESULT 7

```
US-10-012-231A-274
; Sequence 274, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-231A-274
```

```
Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1
```

US-10-803-530-2 (1-435) x US-10-012-231A-274 (1-2063)

```
QY 2 AppProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCTTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTCGCGCAACCCCGT 278
```

```
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTAGACCTG 338
QY 42 AlaSerIleIleLeValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 339 GCGAGTATCATCTTGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGGGCAGCCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGAGAGCTGAGACTGT 458
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGGAGGAGGAGGACCTGTCTCAAGAGCTTCCCGAAGGGCTCGAGTGGCA 518
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCACAGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGGCTGTTCGACAACTTCACGAAGCTCTCGCTGACAGACGCTGTAGGCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCGACAGACGAGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATCAGACAAAACAGCCAGGAGCTTCGCATGCGGAATCTCAAGTGGGCGCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGCTCTCCCTGCTGCTGTGCTGTGGAGAGAGCCCTGAAGACCCCC 803
QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 COTGTGGTGGGTGGGGAGGAGGCTCTGTGGATTCTTGCCCTTGGCAGGTCAAGTCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TAGGACAAACAGCAGCTCTGTGGAGGAGCATCTTGGACCCCTCTGGTCTCTCAGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGCGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGACGTCTCCACTCTCTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACATACTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGCAGGCGTCAGTCCAGGTCAATCCAGCACACACCGTCAATGCAAGACGATCGTACCA 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCAACGAGAGAGATGATGTGTGACAGGCATCCCGAAGGGGGTGTGACACCTGC 1343
```

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGCCCTGATGATACCAATCTGACAGTGGCATGTGGTGGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGACCCAGAGGTATACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGAATCTACATATGCTGGAAGGCTGAGCTG 1505

RESULT 8

US-10-015-389A-274
; Sequence 274, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-389A-274

Alignment Scores:
Pred. No.: 1.67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-389A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACCTCTGAAAGAGCTCGATGTCAAAACCCCTGCGCAACCCCGT 278
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCCATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTAGCCCTG 338
QY 42 AlaSerIleIleIleValValValIleIleValIleValIleValIleValIleValIle 61
Db 339 GCGAGTATCATATGTGGTGTGCTCATCAAGGTGATCTGATAAATACTACTTCTCTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TGGGGGAGGACCTCTCCACTTTCATCCCGAGAGCAGCTGTGTGACCGAGAGCTGGACTGT 458
QY 82 ProLeuGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAGGGCTCGAGTGGCA 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCCCTCTCAAGGACCGATCCACACTGACAGGTGTGGACTCGGCACACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysAspGlnMet 141
Db 579 TTCTCTGCCCTGTTCGAACTTCACAGAAAGCTCTCGCTGAGACAGCGCTGTAGSCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCAGACCCAGGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATTCAGAAAAACAGCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTGTGGATCTTTGGCTTGGCAGGTGACATCCAG 803
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGGAGGCTCTGTGTGGATCTTTGGCTTGGCAGGTGACATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TAGCAGAAAAACAGCAGCAGCTCTGTGGAGGAGCATCTCGACCCCTCTCTCACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAACAATACCGATGTGTCACTGGAAGGTGGGAGCAGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGCAGCTTCCCATCCCTGGCTGTGGCCAGATCATCATCATTTGAATTCACACCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGACGTCCCATCTCACTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAGATGAGGGAAGATGTCTGACATCTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGAGGGCTGATTCAGGTCAATGACAGCACAGGTGCAATGCAATGCAATGCAATGCAATG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValValAspThrCys 381
Db 1284 GGGGAAGTCACGAGAAGATGATGTGTGACGGCATCCCGGAAGGGGGTGTGGACACCTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGTGTACAGTGGTGGGCCCTCATGTACCAATCTGACCAAGTGGCATGTGTGGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCCGAGACCCAGAGGATATACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGAATCTACATGTCTGGAAGGCTGAGCTG 1505

RESULT 9

US-10-006-768A-274
; Sequence 274, Application US/10006768A
; Patent No. 6936697


```
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830PIC47
; CURRENT APPLICATION NUMBER: US/10/015,671A
; CURRENT FILING DATE: 2001-12-11
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-671A-274

Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservativeness: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-671A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACTCTGACAGCCTCGATGTCACACCCCTCGCGCAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyLeuProIleIleAlaLeuSerLeu 41
Db 279 ATCCCATGGAGACCTTCAGAAAGGTGGGATCCCATCATAGCACTACTGAGCCTG 338

QY 42 AlaSerIleIleValValValLeuIleLysValIleLeuAspLysTyrrPheLeu 61
Db 339 GCGAGTATCATCATGTTGTTCTCATCAAGGTGATTCAGTAAATCTACTTCCTC 398

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGGGCAGCCTCTCCACTTCATCCGAGGAGCAGCTGTGTGCGGAGAGCTGGACTGT 459

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGAGCAGAGCAGTGTGTCAAGAGCTTCCCGAAGGGCTGCGAGTGCA 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAGGACCGATCCACACTGCGAGGTGCTGGACTCGGCCACAGGAACTGG 578

QY 122 PheSerAlaCysPheAspAspPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCGACAACTTCACAGAACTCTCGCTGAGACAGCCTGTAGSCAGATG 638

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGAGATTGGCCAGACCCAGGATCTGGAT 683

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTTGTGAATTCACAGAAAACAGCAGAGCTTCGCATGCGGAATCAAGTGGGCCCTGT 743

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGCTCCCTGGTCTCCCTGCACTGTCTTGGCTGTGGGAGAGCCTGGAAGACCCCC 803

QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGGAGGCTCTGTGGATCTTGGCTTGGCAGGTGAGCATCCAG 863

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACACAAACAGCAGCGTCTGTGAGGAGGAGCATCTCGAAGGAGGAGGAGGAGGAGGAGGAG 923

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCGCACTGCTTCAGGAAACATACCGATGTTCACTGGAAGGTGCGGCGGCTCAGAG 983
```

```
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGAGCTTCCCATCCCTGGCTGGCCCAAGATCATCATCATGATTAATCAACCCC 1043

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGCAGTTCCTCCACTCCTTTCTCA 1103

QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGCTGCTCCCTTCTTTGATGAGGAGCTCATCCAGCCACCCCA 1163

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGCTTTTACGAAGCAGAAATGGAGGGAAGATGCTCTGACATACTG 1223

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGCGAGCGCTCAGTCCAGGTCAITGACAGCACAGGTGCAATGCAGCGATGGGTACAG 1283

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAGTCAACCCGAGAAGATGATGTGTGAGGCGCATCCCGAAGGGGTGTGGACACTGC 1343

QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACAGTGGCATGTGGTGGGCATC 1403

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValThrLysValSer 421
Db 1404 GTTAGCTGGGCTATGGCTCGGGGGCCCGAGCAGCCCCAGGAGGTATACACCAAGGTCTCA 1463

QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505

RESULT 11
US-10-015-393A-274
; Sequence 274, Application US/10015393A
; Patent No. 6951737
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC46
; CURRENT APPLICATION NUMBER: US/10/015,393A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-393A-274

Alignment Scores:
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservativeness: 0
```

Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-015-393A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 219 GATCTGACAGTATCAACCTCTGACAGCCTCGATGTCAACACCCCTGGCAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 279 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCTG 338

QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
DB 339 GCGAGTATCATATTGTGGTTCCTCATCAGAGTGTCTGGATAATACTACTTCTC 398

QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
DB 399 TCGGGGAGCCTCTCCACTTCATCCGAGGAGCAGTGTGTGACGAGAGCTGACTGT 458

QY 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
DB 459 CCTTTGGGGAGGAGCAGAGCACTGTGTCAAGAGCTTCCCGAAGGGCCTGCACTGGCA 518

QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
DB 519 GTCCGCCCTCTCAAGGACGATCCACACTGCAAGTGTGGACTCGGCCACAGGAACTGG 578

QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 579 TTCTCTGCTGTTTCGACACTTCACAGAGCTCTCGTGAGACAGCCTGAGGAGATG 638

QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
DB 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCAGACCCAGGATCTGGAT 683

QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
DB 684 GTTGTGTAATCACAGAAACAGCAGGAGCTTCGCAATCGCGAACTCAAGTGGGCCCTGT 743

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
DB 744 CTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTGTGCTGTGGAGAGCCTGAAGACCC 803

QY 202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 804 CGTGTGGTGGTGGGAGAGGCTCTGTGGATTTCTGGCCTTGGCAGGTACGATCCAG 863

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
DB 864 TAGCAAAACAGCAGCTGTGTGGAGGAGCATCTGGACCCCACTGGTCTCTCAGGCA 923

QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
DB 924 GCCCACTGCTTCAGGAAACATACCATGATGTTCCTCACTGGAAGTGGGGCAGGCTCAGAC 983

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
DB 984 AAACTGGGAGCTTCCCATCTCCCTGGCTGTGGCAAGATCATCATATTCAACCC 1043

QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1044 ATGTACCCCAAGACAAATGACATCGCCCTCATGAAGCTGCAGTCTCCACTTCTCTCA 1103

QY 302 GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321
DB 1104 GGCACAGTCAGGCCCATCTGTCTGGCCTTCTTTGATGAGAGCTCACTCCAGGCCACCA 1163

QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
DB 1164 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGGGAGGAGATGTCTGACATCTG 1223

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1224 CTGCAGGCGTCAAGTCCAGGTCAATTCAGACAGCACACGCGTCAATGCAGACGATCGTACCA 1283

QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
DB 1284 GGGGAAGTCAACGAGAAGATGATGTGCAGGATCCCGAAGGGGTGTGGACACCTGC 1343

QY 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
DB 1344 CAGGTGACAGTGGTGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGTGGGCATC 1403

QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
DB 1404 GTTAGCTGGGCTATGTGTGGGGGCCGAGCACCCAGGAGTATACACCAAGGTCTCA 1463

QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505

RESULT 12

US-10-011-833A-274
; Sequence 274, Application US/10011833A
; Patent No. 6951920

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C22
; CURRENT APPLICATION NUMBER: US/10/011,833A
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-011-833A-274

Alignment Scores:

Pred. No.: 1.67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x US-10-011-833A-274 (1-2063)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 219 GATCTGACAGTATCAACCTCTGACAGCCTCGATGTCAACACCCCTGGCAACCCCGT 278

QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
DB 279 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTACTGAGCCTG 338

QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61


```
Db 339 GCGAGTATCATTTGTTGTTCTCTCATCAAGGTGATTCTGGATAAATACTACTTCTC 398
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGCGGAGCGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGACATGTGTCAAGAGCTTCCCGAAGGGCCCTGCGAGTGGCA 518
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCCCTCTCCAGGACCGATCCACACTGACGTGTGTGACTGGCCACAGGGAATGG 578
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCTGTAGGACATG 638
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCGACAGACCGAGATCTCGAT 683
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGAAATCAGAAAACAGCAGGAGCTTCGATGCGAACTCAAGTGGGCCCTGT 743
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLeuThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCTCACTGTCTTGCCTGTGGGAAGAGCCTGAAGACCCC 803
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGTGGGTGGGAGGAGGCTCTGTGGATTCTTGGCCCTTGGCAGGTGAGCATCCAG 863
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTCTGTGGAGGAGGAGCATCTGGACCCCTCTGGGCTCTCAGCGCA 923
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAACATACCGATGTGTTCACCTGGAAGGTGCGGCGAGGCTCAGAC 983
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 984 AAACCTGGGAGCTTCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACACCC 1043
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATGCCCTCATGAAGCTGCGATTCACCACTCACTTCTCA 1103
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCATCTGTCTGCCCTCTTTTCATGAGGAGCTCACTCCAGCCACCCCA 1163
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATGTGATGGGGCTTTACGAAGCAGAATGGAGGAAAGATGTCTGACATACTG 1223
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1224 CTGACGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCATGACAGCATGCGGTACAG 1283
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGTGACGCATCCCGNAGGGGTGTGGACACCTGC 1343
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTCACAGTGTGGGGCCCTGATGTACCAATCTGACCATGTGGCATGTGGTGGGCATC 1403
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTGGGGGGGCCCGAGACCCCGAGGATATACCAAGGTCTCA 1463
```

```
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505
```

RESULT 13

```
US-10-006-041A-274
; Sequence 274, Application US/10006041A
; Patent No. 6951921
```

```
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
```

```
; APPLICANT: Botstein, David
```

```
; APPLICANT: Desnoyers, Luc
```

```
; APPLICANT: Eaton, Dan I.
```

```
; APPLICANT: Ferrara, Napoleone
```

```
; APPLICANT: Fong, Sherman
```

```
; APPLICANT: Gao, Wei-Qiang
```

```
; APPLICANT: Goddard, Audrey
```

```
; APPLICANT: Godowski, Paul J.
```

```
; APPLICANT: Grimaldi, Christopher J.
```

```
; APPLICANT: Gurney, Austin L.
```

```
; APPLICANT: Hillan, Kenneth J.
```

```
; APPLICANT: Pan, James
```

```
; APPLICANT: Paoni, Nicholas F.
```

```
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
```

```
; FILE OF INVENTION: Acids Encoding the Same
```

```
; FILE REFERENCE: P2830P1C8
```

```
; CURRENT APPLICATION NUMBER: US/10/006,041A
```

```
; CURRENT FILING DATE: 2001-12-06
```

```
; Prior Application removed - See File Wrapper or Palm
```

```
; NUMBER OF SEQ ID NOS: 477
```

```
; SEQ ID NO 274
```

```
; LENGTH: 2063
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
```

```
US-10-006-041A-274
```

Alignment Scores:

```
Pred. No.: 1,67e-232 Length: 2063
Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 3 Gaps: 1
```

```
US-10-803-530-2 (1-435) x US-10-006-041A-274 (1-2063)
```

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
```

```
Db 219 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGCAACCCCTGCGCAACCCCGT 278
```

```
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
```

```
Db 279 ATCCCATGAGACCTTCAGAAAGGTGGGATCCCCATCATATGACACTACTGAGCTG 338
```

```
Qy 42 AlaserIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPheLeu 61
```

```
Db 339 GCGAGTATCATCATTTGTGTTGCTCTCATCAAGTGATTCTGGATAAATACTACTTCTC 398
```

```
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
```

```
Db 399 TCGGGGAGCGCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
```

```
Qy 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
```

```
Db 459 CCCTTGGGGAGGACGAGGAGCAGCTGTGTCAAGAGCTTCCCCAAGGGCCCTGCGAGTGGCA 518
```

```
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
```

```
Db 519 GTCCGCCCTCTCCAAAGGACCGATCCACACTGACGTGTGAGCTCGGCCACAGGGAACCTGG 578
```

```
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
```



```
Db 579 TTCTGCTGCTGTTTCGACAACTTACACAGAGCTCTCGCTGAGACAGCCTGTAGGACAGATG 638
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCCCAGACAGGATCTGGAT 683
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATACACAGAAAAGCAGAGAGCTTCGCATGCGGAATCAAGTGGGCCCTGT 743
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGCTGTGGAGAGAGCTGAAGACCC 803
Qy 202 ArgValValGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGAGGCTCTCTGTGATTCTTGCCCTTGGCAGGTTCAGCATCCAG 863
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTCTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCTACGGCA 923
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTCAGAAACATACCATGATGTTCAACTGGAAGGTGCGGGCAGGCTCAGAC 983
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AACTGGGAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGATTCAATTCACCC 1043
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAAATGACATGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCA 1103
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCAGAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCA 1163
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTGGATGGGGCTTTACGAAGCAGAAATGGAGGGAAGATGTCTGACATCTG 1223
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGAGGGCTGATGTCAGGTCATTGACAGCACAGTGCATGACAGAGATGCGTACCAG 1283
Qy 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAAGATGATGTGACAGCATCCCGAAGGGGTGTGGACACCTGC 1343
Qy 382 GlnGlyAspSerGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGCCCTGATGATACCAATCTGACAGTGGCATGTGGTGGGCATC 1403
Qy 402 ValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTACCTGGGGCTATGGCTGCGGGGGCCGAGCACCCAGGAGTATACACCAAGTCTCA 1463
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCACTGGATCTACAAATGTCTGGAAGGCTGAGCTG 1505
```

RESULT 14

```
US-10-012-064A-274
; Sequence 274. Application US/10012064A
; Patent No. 6953841
```

GENERAL INFORMATION:

```
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
```

```
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C19
; CURRENT APPLICATION NUMBER: US/10/012,064A
; CURRENT FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 274
; LENGTH: 2063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-064A-274
```

Alignment Scores:

Pred. No.:	1.67e-232	Length:	2063
Score:	2297.50	Matches:	429
Percent Similarity:	98.85%	Conservative:	0
Best Local Similarity:	98.85%	Mismatches:	0
Query Match:	98.10%	Indels:	5
DB:	3	Gaps:	1

US-10-803-530-2 (1-435) x US-10-012-064A-274 (1-2063)

```
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCTCTGCGCAACCCCGT 278
Qy 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCATGAGAGACCTTCAGAAAGGTGGGGATCCCCCATCATCATAGCATCTAGCGCTG 338
Qy 42 AlaSerIleIleIleValValValValIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 339 GCGAGTATCATCATTTGTGTTCTCATCAAGGGTATCTTGGATAAATACTACTTCTTC 398
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TCGGGGAGCTCTCCACTTTCATCCGAGGAGAGAGCTGTGTGACGAGAGAGCTGACTGT 458
Qy 82 ProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCCTTGGGGAGGACGAGGAGCACCTGTGTCAAGAGCTTCCCCGAAGGGCTCAGTGGCA 518
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
```

```
Db 519 GTCCGGCTCTCCAGGACCGATCCACACTGCGAGGTCTGGACTCGGCCACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTTTCGACAACTTCACAGNAGCTCTCGCTGACAGAGCTGTAGGAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GGCTACAGC-----AGAGCTGTGGAGATTGGCCACAGACCGAGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATCACAGAAACAGCCAGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCTCGTCTCCCTGACCTGTCTTGCTGTGGAGAGAGCTGAAACACCCC 803
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGTGGTGGGAGAGAGGCTCTCTGTGATTCTTGCCCTTGGCAGGTCCAGATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACGACAAACAGCAGCTGTGTGGAGGGAGCATCTCGGACCCCACTGGGTCTCTCACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGGAACATACCATGTGTTCACCTGGAAAGTGGGGAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAATGGGCACTTCCCATCCCTGGCTGTGGCCAAAGATCATCATTTGAATTCACCCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGAACATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCAAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 1164 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAGATGCTCGACATCTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 1224 CTGAGCGGTGAGTCCAGGTCAATTGACAGACACCGGTGCAATGCAGACGATGGGTACCAG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCACCGAGAGATGATGTGTGAGGAGATCCCGAAGGGGGTGTGGACACTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGGTGACAGTGTGGGGCCCTGATGTACCAATCTGACCACTGGCATGTGTGGTGGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGACCCCGAGAGATATACCAAGGTCTCA 1463
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAAATGTCTGGAAAGGCTGAGCTG 1505
```

RESULT 15

US-09-607-745-8

; Sequence 8, Application US/09607745

; Patent No. 6750034

; GENERAL INFORMATION:

; APPLICANT: Darrow, Andrew L

; APPLICANT: Qi, Jain-shen

```
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
US-09-607-745-8
```

Alignment Scores:

Pred. No.:	3,72e-123	Length:	1189
Score:	1259.00	Matches:	227
Percent Similarity:	97.86%	Conservative:	2
Best Local Similarity:	97.01%	Mismatches:	5
Query Match:	53.76%	Indels:	0
DB:	3	Gaps:	0

US-10-803-530-2 (1-435) x US-09-607-745-8 (1-1189)

```
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 163 AAGATCGTTGGGGCTTATGCTCTAGATGTGGATTCTTGGCTTGGCAGGTCCAGATCCAG 222
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 223 TAGCAAAACAGCACGCTGTGTGGAGGAGCATCTCTGGACCCCACTGGGTCTCTCACGGCA 282
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 283 GCCCACTGCTTCAGGAACATACCGATGTGTTCACCTGGAAGGTGGGCGAGCTCAGAC 342
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 343 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCCAAAGATCATCATCTTGAATTCACCCC 402
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 403 ATGTACCCCAAGAACATGACATCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCA 462
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 463 GGCAAGTCAGGCCCATCTGTCTGCCCTTCTTTGATGAGGAGCTCACTCCAGCCACCCCA 522
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 523 CTCTGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAGATGCTCGACATCTG 582
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
Db 583 CTGCAAGGCTCAGTCCAGGTCAATTGACAGCACCGGTGCAATGCAGACGATGGGTACCTG 642
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 643 GGGGAAGTCACCGAGAGATGATGTGTCCAGGCATCCCGAAGGGGGTGTGGACACCTGC 702
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 703 CAGGGTGACAGTGTGGGGCCCTGATGTACCAATCTGACCACTGGCATGTGTGGTGGGCATC 762
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 763 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGACCCCGAGGGGTATACCAAGGTCTCA 822
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 823 GCCTATCTCAACTGGATCTACAAATGTCTGGAAAGGCTGAGCTG 864
```

Search completed: December 6, 2005, 23:11:01
Job time : 259 secs



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 19:52:16 ; Search time 4188 Seconds
(without alignments)
4859.690 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2342
Sequence: 1 MDPDSQPLNSLDVPLRKP.....VTKVSALNVIYNVWRAEL 435

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO_spool_p/US10803530/runat 05122005 083216 2133/app query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10803530 @CGN 1 1 8010 @runat 05122005 083216 2133 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

AY420489 995 bp DNA linear GSS 17-DEC-2003
Homo sapiens TMRSS4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY420489
VERSION AY420489.1 GI:39776446
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS

1 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.

TITLE

Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302

REFERENCE
AUTHORS

2 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.

TITLE

Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

COMMENT

This sequence as made by sequencing genomic exons and ordering them

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1779	76.0	995	10 AY420489	AY420489 Homo sapi
2	1688.5	71.2	2177	4 AK078890	AK078890 Mus muscu
3	1427	60.9	995	10 AY420490	AY420490 Pan trogl
4	1424	60.8	995	10 AY420491	AY420491 Mus muscu
5	1233	52.6	844	2 BE531100	BE531100 601278466
6	1203	51.4	895	5 BU522841	BU522841 AGENCOURT
7	1158.5	49.5	975	2 BG288427	BG288427 602388091

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

FEATURES		based on alignment.	
source	Location/Qualifiers		
1..995	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
gene	<1..>995		
	/gene="TMPRSS4"		
	/locus_tag="HCM7241"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1..15e-176	Length:	995
Score:	1779.00	Matches:	330
Percent Similarity:	99.70%	Conservative:	0
Best Local Similarity:	99.70%	Mismatches:	1
Query Match:	75.96%	Indels:	0
DB:	10	Gaps:	0
US-10-803-530-2 (1-435) x AY420489 (1-995)			
QY	52	LysValIleLeuAspLysTyrPheLeuCysGlyGlnProLeuHisPheIleProArg	71
Db	3	AAAGTGATTCTGGATAAATACTACTTCTCTCGGGGAGCCCTCTCCATTTATCCCGAGG	62
QY	72	LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGlnAspGluHisCysVal	91
Db	63	AAGCAGCTGTGTGACGAGAGCTGGACTGTCTCTTGGGGAGCAGGAGCAGCTGTGTG	122
QY	92	LysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeu	111
Db	123	AGAGCTTCCCCAAGGGCCTCGAGTGGAGTCCGCTTCCAGAGCCGATCCACACTG	182
QY	112	GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetrGlu	131
Db	183	CAGGTGCTGGACTCGGCCACAGGAACTGTTCTCTGCTGTTTCGACAACTTCACAGAA	242
QY	132	AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAla	151
Db	243	GCTCTCGCTGAGACAGCTGTAGGCAGATGGCTACAGCAGCAAAACCCACTTTCAGAGCT	302
QY	152	ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu	171
Db	303	GTGGAGATTGGCCAGACAGAGATCTGGATGTTGTTGAAATCACAGAAAAACAGCCAGAG	362
QY	172	LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys	191
Db	363	CTTCGATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGT	422
QY	192	LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVal	211
Db	423	CTTGCTGTGGGAAGAGCCTGAAGACCCCGTGTGTGGTGTGTGGAGGCGCTCTGTG	482
QY	212	AspSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySer	231
Db	483	GATCTTGGCTTGGCAGTCCAGATCCAGTACGACAAACAGCAGCTGTGTGGAGGAGC	542
QY	232	IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal	251
Db	543	ATCCTGGACCCCACTTGGGTCTCTCAGCGACCCCACTGCTTCAGAAAACATACCGATGTG	602
QY	252	PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal	271
Db	603	TTCACTGGAAGGTGGCGGAGGCTCAGACAAATCGGGCAGCTTCCCATCCCTGGCTGTG	662
QY	272	AlaLysIleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu	291
Db	663	GCCAAAGATCATCATCATTTGATTCAACCCCATGTACCCCAAGACATGATCATCGCCCTC	722
QY	292	MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe	311
Db	723	ATGAAGCTGCAGTTTCCCACTCACTTCTTCAGGGCAGTCAGGCGCCATCTGTCTGCGCCCTC	782
QY	312	PheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLys	331
Db	783	TTTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGGATGGGTTTACGAAG	842
QY	332	GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer	351
Db	843	CAGAAATGGAGGGAAGATGTTCTGACATACTCTGTCAGGGCTCAGTCCAGGTCAATGACAGC	902
QY	352	ThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAla	371
Db	903	ACACGGTGCAATGCAGACGATCGTACAGGGGAACTCCCGAGAAGATGATGTGTGCA	962
QY	372	GlyIleProGluGlyGlyValAspThrCysGln	382
Db	963	GGCATCCCGAAGGGGTGTGGACACCTGCCAG	995
RESULT 2			
AK078890			
LOCUS			
DEFINITION			
	AK078890	2177 bp	mus musculus adult male colon cDNA, RIKEN full-length enriched library, clone:903062G02 product:SIMILAR TO TRANSMEMBRANE PROTEASE, SERINE 4 homolog [Mus musculus], full insert sequence.
ACCESSION			
VERSION	AK078890	GI:26098158	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci, P. and Havaehizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	10349636		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	11042159		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBMED	11076861		
REFERENCE			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 695-690 (2001)		
REFERENCE			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE			
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,		

Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://phantom.gsc.riken.jp/.

FEATURES

source

```

1..2177
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9030622G02"
/db_xref="taxon:10090"
/clone="9030622G02"
/sex="male"
/tissue_type="colon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
304..1507
/note="SIMILAR TO TRANSMEMBRANE PROTEASE, SERINE 4 homolog [Mus musculus] (SPTR|AAH21368, evidence: FASTY, 92%ID, 100%length, match=1201) putative"
2154..2159
/note="putative"
2177
/note="putative"

```

misc_feature

polyA_signal

polyA_site

ORIGIN

Alignment Scores:

```

Pred. No.:      1,74e-164      Length:      2177
Score:          1668.50      Matches:      310
Percent Similarity: 80.05%      Conservative: 39
Best Local Similarity: 71.10%      Mismatches:  51
Query Match:      71.24%      Indels:      36
DB:               4           Gaps:         1

```

US-10-803-530-2 (1-435) x AK078890 (1-2177)

```

QY      1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspVallysProLeuArgLysPro 20
Db      304 ATGAGTCACAGAGTGGCCCAACCCCTCAACACCGGTATATTGTCCTTCGCCAACCC 363
QY      21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSer 40
Db      364 CGAAGGCCCCAGGAGACCTTCAAA----- 387
QY      41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTyrPhe 60
Db      387 ----- 387
QY      61 LeuCysGlyGln-ProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAs 80
Db      388 -----AAGTCCCTTGACCTTCATTACAGAGGGCCAGTTGTGTGACGGCCACCTTGA 438
QY      80 pCysProLeuGlyCysGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      439 CTGGCCCTCAGGGGAGGATGAGGAACACTGTGTCAAGGACTTCCCTGAAAGCCCGAGT 498

```

```

QY      100 lAlaValArgLeuSerLysAspArgSerThrIleuGlnValLeuAspSerAlaThrGlyAs 120
Db      499 GGCAGTCCGGCTCCCAAGGACAGATCCACCTCGAGTGTGGATGAGCCACAGGGAC 558
QY      120 nTriPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgG1 140
Db      559 CTGGCCCTCAGTCTGTTTCGACAACTTCACAGAGCACTGGCCAAAGACAGCCCTGCAGACA 618
QY      140 nMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGlnAspLe 160
Db      619 GATGGCTATCAGACGCCCGCGCTTTCAGAGCAGTGGAGATCCGTCAGATCAGAACCCT 678
QY      160 uAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPr 180
Db      679 CCCTGTTCTCAAGTCACAGAAACAGCCAGCAACTTCAGGTGCAGAAATGGAAAGAGCTGAAGAC 738
QY      180 oCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysTh 200
Db      739 CTGCTCTCAGGCTCCCTGGTTCTTCCTGGCTGCTTCTGACTGTGGAAAGAGCTGAAGAC 798
QY      200 rProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIl 220
Db      799 TCCTCGTGTGGTGGGTGGAGGCCCTCTGGATTCTTGGCGGTGGCAGTCAAGCAT 858
QY      220 eGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuTh 240
Db      859 CCAGTACAAACAGCAGCATGTCTGTGGTGGGAGCATCTCTGGATCCCACTGGATCTCTAC 918
QY      240 rAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySe 260
Db      919 ACCAGCCCACTCTCTCAGGAAGTATCTGTATGTGTCAAGCTGGAGGTCAAGGCGGCGCTC 978
QY      260 rAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAs 280
Db      979 AAACATACTGGTAACTCTCCATCTTGCCTGTGGCCAAAGATCTTCACTCGCTGAACCCAA 1038
QY      280 nProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPh 300
Db      1039 TCCTCTGTACCCCAAGAGAGAGACATTGCTCCCTTGTAAAGTGCAGATGCCACTCACT 1098
QY      300 eSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaTh 320
Db      1099 CTCAGGTCAGTCAGCCCATCTGCTGCCCTCTCTGATGAGGTCTTGTCCAGCCAC 1158
QY      320 rProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAsp1 340
Db      1159 ACCAGTCTGGTCAATGGATGGGCTTTACAGAGAAACCGGAGGAAAGATGTCTGACAT 1218
QY      340 eLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTy 360
Db      1219 GCTACTGCAGGCATCAGTCCAGTCAATTCACAGCACACCGTGCATTCAGAGATGCCTA 1278
QY      360 rGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspTh 380
Db      1279 CGAAGGGGAAGTCACCGCTGAGATGCTGTGTGAGGTACCCACAGGGTGGCAGGACAC 1338
QY      380 rCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValG1 400
Db      1339 CTGCCAGGTGACAGTGTGGGCTTTGATGTACCACTTCTGACAAGTGGCAGGTAGTAGG 1398
QY      400 yIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVa 420
Db      1399 CATCGTGTGGTGGGCCATGGATGCGCGGCCCAAGTACTCTCTGGAGTGTATACCAAGGT 1458
QY      420 lSerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      1459 CACTGCTTATCTCACTGGATCTACAATGTTCGGAAGTCTGAGATG 1504

```

RESULT 3

AY420490

LOCUS

DEFINITION

Pan troglodytes

TMPRSS4

gene,

VIRTUAL

TRANSCRIPT,

partial

sequence,

GSS

17-DEC-2003

linear

DNA

995 bp


```

genomic survey sequence.
AY420490
AY420490.1 GI:39776447
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Pan.
1 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence is made by sequencing genomic exons and ordering them
based on alignment.
Location/Qualifiers
1..995
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>995
/feature="TMPRSS4"
/locus_tag="HCM7241"

ORIGIN
Alignment Scores:
Pred. No.: 1
Score: 1.7e-139 Length: 995
Percent Similarity: 100.00 Matches: 273
Best Local Similarity: 82.78% Conservative: 1
Query Match: 60.93% Mismatches: 57
DB: 10 Indels: 0
Gaps: 0

US-10-803-530-2 (1-435) x AY420490 (1-995)

Qy 52 LysValIleLeuAspLysTyrPheLeuCysGlyGlnProLeuHisPheIleProArg 71
Db 3 AAGGTGATCTGGATAAATACTACTTCTCGCGGAGCCCTCCACCTTCATCCCGAGG 62

Qy 72 LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVal 91
Db 63 AAGCAGCTGTGTGACGGAGCTGCAGTCTCCCTGGGGAGGACGAGGACCTGTGTC 122

Qy 92 LysSerPheProGluGlyProAlaValAlaValargLeuSerLysAspArgSerThrLeu 111
Db 123 AAGAGCTTCCCGAAGGCGCTGCAGTGCAGNNNNNNNNNNNNNNNNNNNNNNNNNN 182

Qy 112 GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGlu 131
Db 183 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 242

Qy 132 AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAla 151
Db 243 NNTCTCGCTGAGACAGCTGTAGGACAGTGGGTACAGCAGCAAAACCCACTTTCAGAGCT 302

Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 303 GTGGAGATTGGCCAGACCGAGGATCTGGATGTTGTAATATCACAGAAACAGCCAGGAG 362

Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys 191

```

```

Db 363 CTTACATGCGGAACCTCAAGTGGCCCTGTCTCTCAGGCTCCCTGGTCTCCTGCACGTGT 422
Qy 192 LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVal 211
Db 423 CTTGCTGTGGGAGAGCGCTGANNNNCCCCCGTGTGGTGGCGGGAGGAGGCTCTGTGTG 482
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySer 231
Db 483 GATTCTTGGCCCTTGGCAGGTCAAGNATCCAGTACGNNNNNNNANNANGTCTGTGGAGGAGC 542
Qy 232 IleLeuAspPheHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal 251
Db 543 ATCTCGACCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNAACATACCAGATGTG 602
Qy 252 PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal 271
Db 603 TTCAACTGGAAGGTGCGGGCGGCTCAGACAACTGGGCAGCTTCCCATCCTTGGCTGTG 662
Qy 272 AlaLysIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu 291
Db 663 GCCAAGATCATCATCATTAATTCACCCCATGTACCCCAAGACAAATGATGATGCGCCTC 722
Qy 292 MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe 311
Db 723 ATGAAGCTGCAGTTCNNNNNNNNNNNNNNNNNNNNNNNGCACAGTCAGGCCCATCTGTGCCCCTTC 782
Qy 312 PheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLys 331
Db 783 TTTGATGAGGAGCTCACTCCAGCCACCCCATCTGGATCATTTGGATGGGGCTTTACGAAG 842
Qy 332 GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 843 CAGATGGGAGGAGATGCTGATCTGCTCAGCGGTCAGTCAGGTCATTTGACAGC 902
Qy 352 ThrArgCysAsnAlaAspAlaTyrGlnGlyValThrGluLysMetMetCysAla 371
Db 903 ACACGGTGCAATGCAGACGATCGTACCAGGGGGAAGTACCAGGAAGATGATGTGTGCA 962
Qy 372 GlyIleProGluGlyGlyValAspThrCysGln 382
Db 963 GGATCCCCGAGGGGGTGTGGACACCTGCCAG 995

RESULT 4
LOCUS AY420491
DEFINITION Mus musculus TmpRSS4 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY420491
VERSION AY420491.1 GI:39776448
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Mus.
1 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 995)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

```

COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.

FEATURES source 1..995 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>995 /gene="TPRSS4" /locus_tag="HCM7241"

ORIGIN

Alignment Scores: Pred. No.: 3.53e-139 Length: 995 Score: 1424.00 Matches: 260 Percent Similarity: 87.61% Conservative: 30 Best Local Similarity: 78.55% Mismatches: 41 Query Match: 60.80% Indels: 0 DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x AY420491 (1-995)

Qy 52 LysValIleLeuAspLysTyrPheLeuCysGlyGlnProLeuHisPheIleProArg 71
Db 3 AAGGTGATTCGGATAAATACTACTTCACTGCGGCAGTCCCTTATTCAGAGG 62
Qy 72 LysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVal 91
Db 63 GGCCAGTTGTGTGACGGCCACCTTGACTGCGCTCAGGGGAGGATGAGAACACTGTGTC 122
Qy 92 LysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeu 111
Db 123 AAGGACTTCCCTGAAAGCCGGAGTGGCAGTCCGGCTCTCAAGCAGACAGATCCACCTG 182
Qy 112 GlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGlu 131
Db 183 CAGGTGCTGGATGAGCCACACAGGGACCTGGGCTCAGTCTGTTCCGACAACTTCACAGAA 242
Qy 132 AlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAla 151
Db 243 GCATGGGCAACAGACAGCTGCGACAGATGGGCTATGACAGCCAGCCCGCTTTCAGAGCA 302
Qy 152 ValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsnSerGlnGlu 171
Db 303 GTGGAGATCCGTCAGATCAGAACTCCCTGTGTCTCAAGTCACAGGAACAGCCAGGAA 362
Qy 172 LeuArgMetArgAsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCys 191
Db 363 CTTCAAGGTGAGATGGAAGCAGAGATCTGCTCTCAGGCTCCCTGTTTCTTGCACCTGC 422
Qy 192 LeuAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluAlaSerVal 211
Db 423 CTTGACTGTGGAAGAGCCTGAAGACTCTCTGTGTGGTGGGTGGAGGCCCTGTG 482
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySer 231
Db 483 GATCTTGTGGCGTGGCAGATCAGCATCCAGTACACAGCAGATGCTGTGGTGGAGC 542
Qy 232 IleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspVal 251
Db 543 ATCCTGATCCCCACCTGGATCCTTCACAGAGGCCACTGTCTCAGGAAGTATCTTGATGTG 602
Qy 252 PheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaVal 271
Db 603 TCAAGCTGGAAGTCAAGGCGAGGCTCAAAACATACCTGGTAACCTCTCCATCTTGCCTGTG 662
Qy 272 AlaLysIleIleIleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeu 291
Db 663 GCCAAGATCTTCATCGCTCAACCAATCTCTGTACCCCAAGAGAGGACATGCGCTT 722
Qy 292 MetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhe 311
Db 723 GTTAAAGCTGCAGATGCACCTACATTTCTCAGGCTCAGTCAAGGCCCATCTGCCTGCCCTTC 782

Qy 312 PheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTyrGlyPheThrLys 331
Db 783 TCTGATGAGGTGCTTGTCCAGGCACACCATGCTGGGTCTATGGATGGGCTTTTACAGAA 842
Qy 332 GlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 843 GAAACGGAGGAAAGATGCTGACATGCTACTGAGGCATCAGTCCAGGTTCATCGACAGC 902
Qy 352 ThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAla 371
Db 903 ACACGGTCAATGCAGAGGATGCTTACGAAGGGGAAGTACCGCTGAGATGCTGTGTGCA 962
Qy 372 GlyIleProGluGlyGlyValAspThrCysGln 382
Db 963 GGTACCCACAGGGTGGCAAGGACACCTGCCAG 995
RESULT 5
LOCUS BE531100 844 bp mRNA linear EST 09-AUG-2000
DEFINITION 601278466F1 NIH_MGC_39 Homo sapiens CDNA clone IMAGE:3610337 5', mRNA sequence.
ACCESSION BE531100
VERSION BE531100.1 GI:9759745
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 844)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Plate: LLCM266 row: d column: 18
High quality sequence stop: 712.
FEATURES source 1..844
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3610337"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/note="Organ: pancreas; Vector: pOTB7; Site:1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Alignment Scores: Pred. No.: 4.07e-119 Length: 844 Score: 1233.00 Matches: 254 Percent Similarity: 91.49% Conservative: 4 Best Local Similarity: 90.07% Mismatches: 14 Query Match: 52.65% Indels: 15 DB: 2 Gaps: 2
US-10-803-530-2 (1-435) x BE531100 (1-844)

Qy 314 GluGluLeuThrProAlaThrProLeuTrpIleleGlyTrpGlyPheThrLysGlnAsn 333
 Db 503 GAGGTGCTTGTCCAGGCACACAGCTTGGGTCTTGGATGGGCTTTACAGAAGAAAC 562
 Qy 334 GlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArg 353
 Db 563 GGAGGAAGATGTTGTGACATCTCTGAGGACATCAGTCCAGGTCATCGACAGCACACGG 622
 Qy 354 CysAsnAlaAspAspAlaThrGlnGlyValThrGluLysMetMetCysAlaGlyIle 373
 Db 623 TGCATGACAGGATGCTTACAGAGGAGAGTGCACGCTGAGATGCTGTGTGAGGTACC 682
 Qy 374 ProGluGlyGlyValAlaThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSer 393
 Db 683 CCACAGGGTGCCAGGACACCTGCCAGGCTGACAGTGGTGGGCTTTGATGATCACTTCT 742
 Qy 394 AspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThr 413
 Db 743 GACAAGTGGCAGGTAGTAGGATCTGTGAGCTGGGGCCATGGATGGGCGGCCAAGTACT 802
 Qy 414 -ProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyrAsn 429
 Db 803 CCTGGAGTGATATACCAAGGTCACTGCCTATCTCAACTGGATCTACAAT 851

RESULT 7
 BG288427 975 bp mRNA linear EST 21-FEB-2001
 LOCUS 602388091f1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4516990 5',
 DEFINITION mRNA sequence.

ACCESSION BG288427
 VERSION BG288427.1 GI:13043459
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10409 row: e column: 23
 High quality sequence stop: 659.

FEATURES
 Location/Qualifiers
 1..975

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4516990"
 /tissue_type="transitional cell papilloma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 93"
 /notes="Organ: b1adder; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3,72e-111 Length: 975
 Score: 1158.50 Matches: 221
 Percent Similarity: 96.94% Conservative: 1
 Best Local Similarity: 96.51% Mismatches: 4

Query Match: 49.47% Indels: 4
 DB: 2 Gaps: 1

US-10-803-530-2 (1-435) x BG288427 (1-975)

Qy 196 LysSerLeuLysThrProArgValValGlyGlyGluAlaSerValAspSerTrpPro 215
 Db 3 AAGACCTGAAGACCCCTGCTGTGGTGGGAGGAGGCTCTGTGGATCTTGGCCT 62
 Qy 216 TrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspPro 235
 Db 63 TGGCAGGTTCAGATCCAGTACGACAAACAGCACGCTCTGTGGAGGGAGCATCTCGACCCC 122
 Qy 236 HisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLys 255
 Db 123 CACTGGGTCTCTCAGCGGAGCCCTGCTTCAAGAAACATACCGATGTGTTCAACTGGAG 182
 Qy 256 ValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIle 275
 Db 183 GTGGGGGAGGCTCAGACAAACTGGGAGCTTCCATCCCTGGCTGGTGGCCAGATCATC 242
 Qy 276 IleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGln 295
 Db 243 ATCATTTGAATTCACCCCATCTGATCCCAAGACATGACATCGCCTCATGAAGCTGCAG 302
 Qy 296 PheProLeuThrPheSerGlyThrValArgProLysCysLeuProPheAspGluGlu 315
 Db 303 TTCCCACTCACTTCTCAGGCACAGTCAGGCCCATCTGCTGCCCTTCTTTGATGAGGAG 362
 Qy 316 LeuThrProAlaThrProLeuTrpIleleGlyTrpGlyPheThrLysGlnAsnGlyGly 335
 Db 363 CTCACCTCAGGCACGCCCACTCTGGATCATTTGGATGGGCTTTACGAAGCAGAATCGAGG 422
 Qy 336 LysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsn 355
 Db 423 AAGATGCTGACATCTCTGAGGCGTCAGTCCAGGTCATTGACAGACACACGCTGCAAT 482
 Qy 356 AlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGlu 375
 Db 483 CGACAGCATGTGTACAGGGGAGTACCAGAGAGATGATGTGTGAGGATCCCGAA 542
 Qy 376 -GlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspG 395
 Db 543 GGGGGGTGTGGACACCTGCCA-GGTGACAGTGGTGGGCGCTTGTATGATACCAATCTGACCA 601
 Qy 395 nTrpHisValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProG 415
 Db 602 GTGGCATGTGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGCC--GAGACCCAGG 658
 Qy 415 yValTyr-ThrLysValSerAla 422
 Db 659 AGTTTACCACCAAGGTCTCAAGC 681

RESULT 8

BE616663

LOCUS

DEFINITION

BE616663

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

BE616663

BE616663.1

GI:9898262

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 920)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE616663 920 bp mRNA linear EST 24-AUG-2000
 601278696f1 NIH_MGC_39 Homo sapiens cDNA clone IMAGE:3610990 5',
 mRNA sequence.

intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Alignment Scores:
Pred. No.: 1 79e-100 Length: 617
Score: 1054.50 Matches: 199
Percent Similarity: 97.56% Conservativeness: 1
Best Local Similarity: 97.07% Mismatches: 4
Query Match: 45.03% Indels: 1
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x BM795149 (1-617)

QY	76	AspGlyGluLeuAspCysProLeuGlyGluAspGluHisCysVallySerPhePro	95
DB	2	GACGGAGCTGGAGTGTCTCCCTTGGGGAGGACGAGGACATGTGTCAAGAGCTTCCCC	61
QY	96	GluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAsp	115
DB	62	GAAGGGCTGAGTGGCAGTCCGCTCTCAGAGCCGATCCACACTGGAGTGTGGAC	121
QY	116	SerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetheGluAlaLeuAlaGlu	135
DB	122	TCGGCCACAGGAACCTGTTCTCTGCTCTTTCGACAACTTCACAGAAGCTCTCGTGAG	181
QY	136	ThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeGly	155
DB	182	ACAGCCTGTAGGCAGATGGCTACAGCAGCAACCCACTTTCAGAGCTGTGGAGATTGGC	241
QY	156	ProAspGlnAspLeuAspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArg	175
DB	242	CCAGACCAGGATCTGGATGTTGTTGAAATACAGAAACAGCAGGAGCTTCGATCGCG	301
QY	176	AnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly	195
DB	302	AACTCAAGTGGCCCTGTCTCAGGCTCCCTGCTCCCTGCACCTGTCTGTGGATTCTGG	361
QY	196	LysSerLeuLysThrProArgValValGly---GlyGluGluAlaSerValAspSerTrp	214
DB	362	AAGAGCCTGAAGACCCCGCTGTGTGTGGTGGGGAAGGANGCCCTCTGTGGATTCTTGG	421
QY	215	ProTrpGlnValSerileGlnTyrAspLysGlnHisValCysGlyGlySerileLeuAsp	234
DB	422	CCTTGGCAGTCAAGATACAGTACGACAAACAGACAGCTGTGTGGAGGAGCATCTGGAC	481
QY	235	ProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrp	254
DB	482	CCCCACTGGGTCTCAGCGCAGGCCCTCTCTTCAGGAAACATACCGATGTGTTCACCTG	541
QY	255	LysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIle	274
DB	542	AAGGTGGGCGAGGCTCAAAACAACTGGGCGAGCTTCCCATCCCTGTGTGGCCAAGATC	601
QY	275	IleIleIleGluPhe	279
DB	602	ATCATCATTTGAATTC	616

RESULT 10

BUS23218

LOCUS

DEFINITION

BUS23218 956 bp mRNA linear EST 13-SEP-2002
AGENCOURT_10154487 NCI_CGAP_C024 Mus musculus cDNA clone
IMAGE:6529864 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
JOURNAL
COMMENT

BUS23218
BUS23218.1 GI:22833656
EST.
Mus musculus (house mouse)
Mus musculus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 956)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14130 row: k column: 16
High quality sequence stop: 706.
Location/Qualifiers
1. 956
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:6529864"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP C024"
/note="Organ: colon; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

FEATURES
source

ORIGIN

Alignment Scores:
Pred. No.: 7 23e-99 Length: 956
Score: 1042.00 Matches: 213
Percent Similarity: 81.25% Conservativeness: 34
Best Local Similarity: 70.07% Mismatches: 48
Query Match: 44.49% Indels: 9
DB: 5 Gaps: 4

US-10-803-530-2 (1-435) x BUS23218 (1-956)

QY	96	GluGlyProAlaValAlaValArgLeuSerLysAspArgSerThrleuGlnValleuAsp	115
DB	46	GAAGAGCCCGGATGGCAGTCCGGCTCTCAAGGACAGATCCACCTGCAGGTGTGGAT	105
QY	116	SerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsnPhetheGluAlaLeuAlaGlu	135
DB	106	GCAGCCACAGGAGCTGGCCCTCAGCTGTGTTTCGACAACTTCACAGACGATCCGTC	165
QY	136	ThrAlaCysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeGly	155
DB	166	ACAGCTGCAGACAGATGGCTATGACAGCCAGCCGCTTTCAGAGCAGTGGAGATCCGT	225
QY	156	ProAspGlnAspLeuAspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArg	175
DB	226	CCAGATCAGAACTCCCTGCTCAAGTCACAGGAAACAGCCAGGAATTCAGGTGCAG	285
QY	176	AnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly	195
DB	286	AATGGAGCAGATCTGCTCTCAGGCTCCCTGGTTCCTTGGCTGCCTTGAATCTGGA	345
QY	196	LysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpPro	215
DB	346	AAGAGCCTGAAGACTCTCTGTGTGTGTGTGGTGGGAGGAGCCCTCTGGATTCTTGGCCG	405
QY	216	TrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspPro	235


```

Db      406 TGGCAGGTCAGATCCAGTACCAAGCAGCATGCTGTGGTGGAGCATCTGGATCCC 465
Qy      236 HsTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLys 255
Db      466 CACTGGATCCTCACAGCAGCCACTGCTTTCAGGAGTATCTTGATGTCAAGCTGGAAG 525
Qy      256 ValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIle 275
Db      526 GTACGGCAGGCTCAACATCTGGTAACTCTCCATCTTGCCTGTGGCCAAGATCTTC 595
Qy      276 IleIleGluPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGln 295
Db      596 ATCGTGAACCATCTCTGTATCCCAAGAGAGACATGCCCTTGTGTAAGCTGCAG 645
Qy      296 PheProLeuThrPheSerGlyThrValArgProIleCysLeuProPhePheAspGlu 315
Db      646 ATGCCACTCACATCTCAGGCTCAGTCAGGCCCATCTGCTGCTTCTGTGATGAGGTG 705
Qy      316 LeuThrProAlaThrProLeuTrpIleGlyTrp-GlyPheThrLysGlnAsn-GlyG 335
Db      706 CTGTGTCACGCCACACCATGCTGGGTCAATGGATGGGGCTTTACAGAGAAACCGGAG 765
Qy      335 LysMetSerAspIleLeuLeuGln-AlaSerVal-GlnValIleAspSerThrArgCy 354
Db      766 GAAGATGCTCTGACATCTACTCAAGGCATCAGTCCAGAGTCATCGAACGACACGGTG 825
Qy      354 s-AsnAlaAspAspAlaTyr---GlnGlyGluValThrGluLysMetMetCysAlaGlyI 373
Db      826 GCATGACAGAGGATGCTACCCAAAGGGGGAAGGACCGCTCGAATGGTGGGGGACAGG 885
Qy      373 lePro---GluGlyGlyValAspThr---CysGlnGlyAspSer---GlyGlyProLeuM 390
Db      886 TACCCACACAGGTGGGCAAGGACACCTGCCAGGGGGGCCAATGGGGGGGCCCTTT 945
Qy      390 et 390
Db      946 TG 947

CV570449 602 bp mRNA linear EST 22-OCT-2004
oe01c10.y1 Human keratoconus cornea, unamplified, (od/oe) Homo
ACCESSION CV570449
VERSION CV570449.1 GI:54470982
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 602)
Rabinowitz, Y., Dong, L. and Wistow, G.
Expressed sequence tag analysis of human keratoconus cornea
JOURNAL Unpublished (2004)
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 01 row: c column: 10
Seq primer: M13Rp1 reverse primer (ABI).
Location/Qualifiers
1. 602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="oe01c10"
/tissue_type="Cornea"
/dev_stage="Adult"

FEATURES
source
RESULT 11
CV570449
LOCUS
DEFINITION
oe01c10.y1 Human keratoconus cornea, unamplified, (od/oe) Homo
ACCESSION CV570449
VERSION CV570449
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 602)
Rabinowitz, Y., Dong, L. and Wistow, G.
Expressed sequence tag analysis of human keratoconus cornea
JOURNAL Unpublished (2004)
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 01 row: c column: 10
Seq primer: M13Rp1 reverse primer (ABI).
Location/Qualifiers
1. 602
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="oe01c10"
/tissue_type="Cornea"
/dev_stage="Adult"

/lab_host="EMDHI08"
/clone_lib="Human keratoconus cornea, unamplified,
(od/oe)"
/note="Organ: Eye; Vector: pCMVSPORT6; Approximately 40ug
total RNA was extracted from 7 adult human keratoconus
corneas. A directionally cloned cDNA library in the
pSPORT1 vector (Invitrogen) was constructed at Bioserve
Biotechnology (Laurel MD) essentially following the
protocols of the SuperScript plasmid system full details
of which are contained in the manufacturer's instruction
manual (http://www.lifetech.com/). First strand synthesis
was carried out using a Not I primer-adaptor
[5'-pGACTAGTCTAGATCGAGCGGCCCT(T)15-3']. cDNA was
cloned in Not I/Sal I sites. EST analysis was performed on
the unamplified library at the NIH Intramural Sequencing
Center (NISC)."
```

ORIGIN

Alignment Scores:

```

Pred. No.: 1.22e-98 Length: 602
Score: 1037.00 Matches: 193
Percent Similarity: 98.97% Conservative: 0
Best Local Similarity: 98.97% Mismatches: 0
Query Match: 44.28% Indels: 2
DB: 7 Gaps: 0
```

US-10-803-530-2 (1-435) x CV570449 (1-602)

```

Qy      243 HsCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLys 262
Db      1 CACTGCTTCAGGAAACATACCGATGTGTTCACTGGAGGTGGGCGAGGCTCAGACAAA 60
Qy      263 LeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleIleIleIleIle 282
Db      61 CTGGGCGAGCTTCCCATCTCCCTGCTGTGGCCAAGATCATCATTTCAATTCACCCCATG 120
Qy      283 TyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGly 302
Db      121 TACCCCAAGACATGACATCGCCCTCATGAAGCTGCAGTTGCCACTCACTTTCTCAGGC 180
Qy      303 ThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeu 322
Db      181 ACAGTCAGGCCCATCTGTCTGCCCTTCTTGATGAGGAGCTCACTCCAGACCCACCCATC 240
Qy      323 TrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeu 342
Db      241 TGGATCATTTGGATGGGGCTTTACGAAGCAGAAATGGAGGAAGATGTTCTGACATCTG 300
Qy      343 GlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGly 362
Db      301 CAGGGCTCAGTCCAGGTCAATTGACAGCACACGGTGCAATGCAGACGATGCGTACAGGG 360
Qy      363 GluValThrGluLysMetMetCysAla-GlyIleProGlu-GlyGlyValAspThrCysG 382
Db      361 GAAGTCACGAGAGAGATGATGTGCAGGGGCATCCGGAAGGGGGGTGTGGACACCTGCC 420
Qy      382 InGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIleV 402
Db      421 AGGGTGACAGTGGTGGGCCCTGATGTACCAATCTGACCAAGTGGCATGTGGTGGGCATCG 480
Qy      402 alSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerA 422
Db      481 TTAGTGGGGCTATGGCTGGGGGGCCGAGCACCAGGAGTATACACCAAGGTCTCAG 540
Qy      422 laTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db      541 CCTATCTCACTGGATCTACATGTCTGGAAGGCTGAGCTG 581
```

RESULT 12

BP253625

BP253625 Sugano cDNA library, kidney epithelial cell Homo sapiens
 cDNA clone HRC04440, mRNA sequence.


```
ACCESSION BP253625
VERSION BP253625.1 GI:52135906
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
TITLE Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
PUBMED 15342556
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.
FEATURES
source
1..592
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC04440"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cDNA library, kidney epithelial cell"
ORIGIN
Alignment Scores:
Pred. No.: 3,49e-97 Length: 592
Score: 1023.00 Matches: 193
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.68% Indels: 0
DB: 3 Gaps: 0
US-10-803-530-2 (1-435) x BP253625 (1-592)
QY 38 LeuLeuSerLeuAlaSerIlelleValValValLeuIleLysValIleLeuAspLys 57
Db 3 CTACTGAGCTGGCGAGTATCATCTGTTGTTCTCATCAGGTGATCTGGATAA 62
QY 58 TyrTyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGly 77
Db 63 TACTACTTCTTGGCGGCGAGCTCTCCACTTCATCCGAGGAAGCAGCTGTGTGACGA 122
QY 78 GluLeuAspCysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGly 97
Db 123 GAGCTGGAAGTGTCCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCGGAAGG 182
QY 98 ProAlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAla 117
Db 183 CTGTGAGTGGCAGTCCGCTCTCAAGGACCGATCCACTGCGAGTGTGGACTCGGCC 242
QY 118 ThrGlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAla 137
Db 243 ACAGGAACATGGTTCCTGCTGCTGTTGCACTTTCACAGAGCTCTCGCTGAGACGCC 302
QY 138 CysArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAsp 157
Db 303 TGTAGGCAGATGGCTACAGCAGCAAAACCCACTTTCAGAGCTGTGGAGATTGCCCGAC 362
QY 158 GlnAspLeuAspValValGluIleThrGluAsnSerGlnLeuLeuArgMetArgAnSer 177
Db 363 CAGGATCTGGATGTGTGTAATACAGAAACAGCCAGGAGCTTCGCATGCGGAATCA 422
QY 178 SerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSer 197
Db 423 AGTGGGCCCCGTCTCTCAGGCTCCCTGGTCTCCCTGCACGTGTCTTGGCTGTGGGAAGGC 482
QY 198 LeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGln 217
Db 483 CTGAAGACCCCCCTGTGTGGTGGGGAGGAGGAGCCCTCTGTGGATTCTTGGCCTTGGCAG 542
QY 218 ValSerIleGlnTyrAspLysGlnHisValCysGlyGly 230
Db 543 GTCAGCATCCAGTACGACAAACAGCAGCTCTGTGGAGG 581
RESULT 13
CV028368 591 bp mRNA linear EST 20-AUG-2004
LOCUS 6865 Full Length cDNA from the Mammalian Gene Collection Homo
DEFINITION sapiens cDNA 5' similar to BC011703, mRNA sequence.
ACCESSION CV028368
VERSION CV028368.1 GI:51486447
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,
Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,
Clingingsmith, T.R., Hartley, J.L., Eposito, D., Cheo, D., Moore, T.,
Simmmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,
Vandenhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.
TITLE Human ORFome Version 1.1: a Platform for Reverse Proteomics
JOURNAL Genome Res. (2004) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc_Vidal@dfci.harvard.edu
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGTTACAGGATCTCTGACAGTGATCAA
BACKWARD: TACAGCTCAGCTTCCAGACATT
Insert Length: 591 Std Error: 7.00
Plate: 11036 row: 09 column: G
Seq primer: ACTGGCGTGTGTTTACAACTGCTGAGCTGGGAAAC
High quality sequence start: 100
High quality sequence stop: 590
POLYA=No.
FEATURES
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/note="Vector: mixed; The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
Donor vector. Reference : MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"
ORIGIN
Alignment Scores:
Pred. No.: 2,49e-96 Length: 591
Score: 1015.00 Matches: 194
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 43.34% Indels: 0
DB: 7 Gaps: 0
```


PUBMED
15342556

COMMENT
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES
source

Location/Qualifiers
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HRC04627"
/tissue_type="kidney"
/cell_type="epithelial cell"
/clone_lib="Sugano cdna library, kidney epithelial cell"

ORIGIN

Alignment Scores:
Pred. No.: 1.93e-95 Length: 583
Score: 1006.50 Matches: 194
Percent Similarity: 97.49% Conservative: 0
Best Local Similarity: 97.49% Mismatches: 0
Query Match: 42.98% Indels: 5
DB: 3 Gaps: 1

US-10-803-530-2 (1-435) x BP253686 (1-583)

Qy	19	LysProArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeu	38
Db	2	AAACCCCGATATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATATAGCACTA	61
Qy	39	LeuSerLeuAlaSerIleIleValValLeuIleLysValIleLeuAspLysTyr	58
Db	62	CTGAGCCTGGCGAGTATCATTTGTTGTCTCATCAAGGTGATCTGGATAAATAC	121
Qy	59	TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu	78
Db	122	TACTTCTCTGGCGGAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTGACGGAGAG	181
Qy	79	LeuAspCysProLeuGluGluAspGluGluHisCysValLysSerPheProGluGlyPro	98
Db	182	CTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTCTAAGAGCTTCCCGAAGGGCCT	241
Qy	99	AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr	118
Db	242	GCAGTGGCAGTCCGCTCTCCAGAGCCGATCCACACTGCAGGTGCTGGACTCGGCCACA	301
Qy	119	GlyAsnTrpPheSerAlaCysPheAspAsnPhThrGluAlaLeuAlaGluThrAlaCys	138
Db	302	GGGAACCTGGTTCTCTGCTGTTTCGACAACTTCACAGAAGCTCTCGCTGAGACAGCCTGT	361
Qy	139	ArgGlnMetClyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGln	158
Db	362	AGGCAGATGGGCTACAGC-----AGAGCTGTGAGATGGCCAGACCAG	406
Qy	159	AspLeuAspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSer	178
Db	407	GATCTGGATGTTGTTGAAATCACAGAAAACAGCCAGGAGCTTCGCATGCCGAACTCAAGT	466
Qy	179	GlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeu	198
Db	467	GGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACGTCTTGCCTGTGGGAAGAGCCTG	526
Qy	199	LysThrProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGln	217
Db	527	AAGACCCCGCTGTGGTGGGTGGAGGAGGCGCTCTGTGATTCTTGGCCTTGGCAG	583

Search completed: December 6, 2005, 23:06:42
Job time : 4200 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 23:06:51 ; Search time 238 Seconds

(without alignments)
568.827 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVPLRKP.....VYTKVSAYLNWYVWKAEL 435

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3289935 seqs, 155610033 residues

Total number of hits satisfying chosen parameters: 6579870

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10803530/runat 05122005 083218 2272/app query.fasta_1.583
-DB=Published Applications NA.New -OFMT=fastap -SUFFIX=rnbpn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -MMAP=US10803530 @CGN 1 1 184 @runat 05122005 083218 2272
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.New:

- 1: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 2: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 3: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 4: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 5: /cgn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 6: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 7: /cgn2.6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 8: /cgn2.6/ptodata/1/pubpna/US11_NEW_PUB.seq2.*
- 9: /cgn2.6/ptodata/1/pubpna/US11_NEW_PUB.seq3.*
- 10: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2297.5	98.1	2063	US-11-102-240-111	Sequence 111, App
2	584.5	25.0	1783	US-11-112-908-14	Sequence 14, Appl
3	580	24.8	1615	US-11-182-752-1	Sequence 1, Appli
4	488.5	20.9	4804	US-11-067-811-3	Sequence 3, Appli
5	475	20.3	2103	US-11-102-240-105	Sequence 105, App
6	473.5	20.2	918	US-11-137-465-23	Sequence 23, Appl
7	467.5	20.0	1129	US-10-131-826A-221	Sequence 221, App
8	449	19.2	1802	US-10-821-234-191	Sequence 191, App

Alignment Scores: 1.13e-243 Length: 2063

Pred. No.: 1.13e-243

2063

9	442.5	18.9	780	7	US-11-056-621-5	Sequence 5, Appli
10	441.5	18.9	2376	6	US-11-056-621-3	Sequence 3, Appli
11	377.5	16.1	1204	6	US-10-131-826A-505	Sequence 505, App
12	369.5	15.8	2016	7	US-11-029-003-7	Sequence 7, Appli
13	356	15.2	1475	6	US-10-623-155-122	Sequence 122, App
14	356	15.2	2294	6	US-10-623-155-123	Sequence 123, App
15	355.5	15.2	2091	7	US-11-029-003-9	Sequence 9, Appli
16	346.5	14.8	999	6	US-10-131-826A-395	Sequence 395, App
17	319	13.6	1006	6	US-10-821-234-422	Sequence 422, App
18	313	13.4	711	6	US-10-401-386B-23	Sequence 23, Appl
19	307.5	13.1	2067	6	US-10-821-234-681	Sequence 681, App
20	306	13.1	1570	6	US-10-131-826A-455	Sequence 455, App
21	305.5	13.0	7681	7	US-11-094-484-2	Sequence 2, Appli
22	303	12.9	711	6	US-10-401-386B-5	Sequence 5, Appli
23	299.5	12.8	2187	7	US-11-094-484-1	Sequence 1, Appli
24	299.5	12.8	5181	7	US-11-094-484-3	Sequence 3, Appli
25	297	12.7	2118	6	US-10-821-234-680	Sequence 680, App
26	291	12.4	768	7	US-11-147-047-10	Sequence 10, Appl
27	289	12.3	1394	6	US-10-401-386B-3	Sequence 3, Appli
28	287.5	12.3	967	6	US-10-401-386B-1	Sequence 1, Appli
29	287.5	12.3	4775	6	US-10-401-386B-62	Sequence 62, Appl
30	285.5	12.2	585	6	US-10-401-386B-11	Sequence 11, Appl
31	284	12.1	1729	6	US-10-401-386B-4	Sequence 4, Appli
32	256.5	11.0	596	7	US-11-108-173-931	Sequence 931, App
33	239	10.2	150468	7	US-11-112-908-56	Sequence 56, Appl
34	239	10.2	193789	7	US-11-112-908-55	Sequence 55, Appl
35	231.5	9.9	906	7	US-11-147-047-11	Sequence 11, Appl
36	230	9.8	2501	6	US-10-821-234-182	Sequence 182, App
37	226	9.6	2846	7	US-11-103-240-37	Sequence 37, Appl
38	217	9.3	597	6	US-10-401-386B-28	Sequence 28, Appl
39	213.5	9.1	21784	7	US-11-182-752-3	Sequence 3, Appli
40	208	8.9	726	7	US-11-137-465-5	Sequence 5, Appli
41	208	8.9	732	6	US-11-137-465-6	Sequence 6, Appli
42	208	8.9	1123	6	US-10-131-826A-247	Sequence 247, App
43	204.5	8.7	1654	6	US-10-401-386B-7	Sequence 7, Appli
44	184.5	7.9	601	7	US-11-182-752-14	Sequence 14, Appl
45	173.5	7.4	1496	6	US-10-750-185-47031	Sequence 47031, A

ALIGNMENTS

RESULT 1

US-11-102-240-111
; Sequence 111, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRES
; FILE REFERENCE: P3230RIC106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 111
; TYPE: DNA
; ORGANISM: Homo Sapien
US-11-102-240-111

Alignment Scores:

1.13e-243 Length: 2063

Score: 2297.50 Matches: 429
Percent Similarity: 98.85% Conservative: 0
Best Local Similarity: 98.85% Mismatches: 0
Query Match: 98.10% Indels: 5
DB: 7 Gaps: 1

US-10-803-530-2 (1-435) x US-11-102-240-111 (1-2063)

```
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db 219 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGAATGTCAAACCCCTGGCGAAACCCCGT 278
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeuLeuSerLeu 41
Db 279 ATCCCCATGAGACCTTCAGAAAGGTGGGATCCCCATCATAGCACTACTGAGCCTG 338
QY 42 AlaserIleIleValValValLeuIleValLysValLysValLysValLysValLys 61
Db 339 GCGAGTATCATNTTGTGTGTCTCATCAAGGTGATTTCTGGATAAATACTACTTCTCTC 398
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 399 TGGGGCAGCCTCTCCACTTCATCCGAGAGAGCAGCTGTGTGACGGAGAGCTGGACTGT 458
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 459 CCTTGGGGAGGAGCAGAGACACTGTGTCAAGAGCTTCCCCGAAGGCCCTGCAGTGCA 518
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 519 GTCCGCTCTCCAGGACCGATCCACACTGCGAGGTGTGGACTCGCCACAGGGAACCTGG 578
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 579 TTCTCTGCTGTGTTCGACAACTTCACAGAGCTCTCGTGAGACAGCTGTAGCAGATG 638
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 639 GCGTACAGC-----AGAGCTGTGGAGATTGGCCCCCAGACCAAGATCTGGAT 683
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 684 GTTGTGTAATCACAGAAACAGCCAGGAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGT 743
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 744 CTCTCAGGCTCCCTGGTCTCCCTGCACCTGTCTGCTGTGGAAAGAGCCTGAAGACCCCC 803
QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 804 CGTGTGGTGGTGGGAGGAGCCTCTGTGGATTCTTGGCCTTGGCAGGTGAGCATCCAG 863
QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 864 TACACAAACAGCAGCTGTGTGGAGGAGCATCTCGACCCCACTGGTCTCTCACGGCA 923
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 924 GCCCACTGCTTCAGAAACATACCAGATGTGTCAACTGGAAGGTGGCGGAGGCTCAGAC 983
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 984 AAATGGGACAGTCTCCATCCCTGGCTGTGGCCAGAGATCATCATTTGAATTCACACCCC 1043
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1044 ATGTACCCCAAGACAATGACATCGCCCTCATGAAGCTGAGTTCACACTCTCTCTCA 1103
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1104 GGCACAGTCAGGCCCACTCTCTGCTGCTCTTTGTATGAGGAGCTCACTCCAGCCACCCCA 1163
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
```

```
Db 1164 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGAATGGAGGAAGATGCTCTGACACTG 1223
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1224 CTGCAGGCGTCAGTCCAGGTCAITGACAGACACGGTGCATATGCAGACGATGCGTACCG 1283
QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1284 GGGGAAGTCAACCGAAGAAGATGATGTGTGAGGATCCCGGAAGGGGTGTGGACACTGC 1343
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1344 CAGGTGACAGTGTGTGGCCCCCTGATGTACCAATCTGACCAGTGGCATGTGTGGGCATC 1403
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1404 GTTAGCTGGGCTATGGCTCGGGGGGCCGAGCACCCAGAGATATACCAAGGTCTCA 1463
QY 422 AlaTyrIleAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1464 GCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG 1505
```

RESULT 2

US-11-112-908-14
; Sequence 14, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Lisa M.
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 14
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-14

Alignment Scores:

Pred. No.:	3-81e-54	Length:	1783
Score:	584.50	Matches:	140
Percent Similarity:	45.0%	Conservative:	60
Best Local Similarity:	31.53%	Mismatches:	161
Query Match:	24.96%	Indels:	83
DB:	7	Gaps:	11

US-10-803-530-2 (1-435) x US-11-112-908-14 (1-1783)

```
QY 19 LysProArgIleProMetGluThrPheArgLysValGlyIleProIleIleAlaLeu 38
Db 288 AGACCAAGTGGGAGCTCTCACT-----GGGGGACCCTGTCTACTTCTGACAGCC 338
QY 39 LeuSerLeuAlaSerIleIleIleValValLeuIleLysValLysValLysValLys 58
Db 339 ATCGGGGCGGATCTCTGGCCATTTGTGCTGTCTCTCAGG----- 380
QY 59 TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu 78
Db 380 ----- 380
QY 79 LeuAspCysProLeuGlyGluAspGluHisCysValLysSerPheProGluGlyPro 98
```

Db 381 -----AGTACCAGGAG-----CCG 395
QY 99 AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr 118
Db 396 CTGTACCCAGTGCAGTCTGCGGACGCTCGGCTCATGCTCTTTGACAAGCGGAA 455
QY 119 GlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCys 138
Db 456 GGGACGTGGCGCTGCTGTGCTCTCGCGCTCAACGCCAGGGTAGCGGATCTACGCTGC 515
QY 139 ArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGln 158
Db 516 GAGGAGATGGCTTC-----CTCAGGGCACTG-----ACCCACTCC 551
QY 159 AspLeuAspValValGluLeuThrGluAsn----- 168
Db 552 GAGCTGGACGTGCAACCGCGCGCGCAATGGCAGCTCGGGCTTCTCTGTGTGGACGAG 611
QY 169 -----SerGlnGluLeuArgMetArgAsnSerSerGlyProCysLeuSer 183
Db 612 GGGAGGCTCCCCACACGAGGCTGTGGAGGTATCTCGGTGTGATTGCCCCAGA 671
QY 184 GlySerLeuValSerLeuHisCysLeuAlaCysGly---LysSerLeuLysThrProArg 202
Db 672 GCGCGTTCTTGGCGCCATCTGCCAGACTGTGGCGCGAGAAAGTGCCTGTGACCGC 731
QY 203 ValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTyr 222
Db 732 ATCTGTGGAGGCGCGGACACAGCTTGGCGCGGTGGCGTGCAGTCAAGCTTCGCTAT 791
QY 223 AspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAla 242
Db 792 GATGAGACACCTCTGTGGGGATCCCTGCTCTCGGGGACTGGGTGTGACAGCGCC 851
QY 243 HisCysPheArgLysHisThrAspValPheAsn---TrpLysValArgAlaGlySerAsp 261
Db 852 CACTGCTTCCGAGCGGACCGGTCCTGTCCGATGGCGAGTGTTCGCGGTCCGCG 911
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 912 GCCAGGCTCTCCACCGTCTGCAGCTGGGGTGCAGGCTGTGCTTACCACGGGGC 971
QY 282 MetTyrPro-----LysAspAsnAspIleAlaLeuMetLys 293
Db 972 TATCTTCCCTTTTCGGACCCCAACAGCAGGAGAACAGCAAGATATTCCTGTGTCAC 1031
QY 294 LeuGlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPheAsp 313
Db 1032 CTCTCAGTCCCTGCCCCCTCACAGAAATACATCCAGCTGTGTGCTCCAGCTCGCGC 1091
QY 314 GluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsn 333
Db 1092 CAGGCGCTGTGGATGGCAAGATCTACCGTAGCGGGTGGGCAACAG---CAGTAC 1148
QY 334 GlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIleAspSerThrArg 353
Db 1149 TATGCCAACAGCGCGGGTACTCCAGAGGCTCGAGTCCCATATCAGCAATGATGTC 1208
QY 354 CysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMetCysAlaGlyIle 373
Db 1209 TGCATATGGCGCTGACTTCTATGGAAACAGATCAAGCCCAAGATGTCTGTGCTGCTAC 1268
QY 374 ProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGln--- 392
Db 1269 CCGAGGGTGGCATTCCTCCAGCGCGACAGCGGTGGTCCCTTTGTGTGTGGAGAC 1328
QY 393 -----SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyrGlyCys 408
Db 1329 AGCATCTCTCGGACGCCACGTTGGCGCTGTGGCATGTGTGAGTTGGGGCACTGGCTGT 1388
QY 409 GlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyr 428

Db 1389 GCCCTGGCCAGAGCCAGGCGTCTACACCAAGTCAGTGACTTCGGGAGTGGATCTTC 1448
QY 429 AsnValTrpLys 432
Db 1449 CAGGCCATAAAG 1460
RESULT 3
US-11-182-752-1
; Sequence 1, Application US/11182752
; Publication NO. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; ORGANISM: HUMAN
US-11-182-752-1
Alignment Scores:
Pred. No.: 1,048-53 Length: 1615
Score: 580.00 Matches: 132
Percent Similarity: 44.06% Conservative: 57
Best Local Similarity: 30.77% Mismatches: 146
Query Match: 24.77% Indels: 94
DB: 7 Gaps: 9
US-10-803-530-2 (1-435) x US-11-182-752-1 (1-1615)
QY 19 LysProArgIleProMetClnThrPheArgLysValGlyIleProIleIleAlaLeu 38
Db 220 AGACCAAGGTGGCAGCTCTCACT-----GGGGGAGCCCTGCTACTTCTGACAGCC 270
QY 39 LeuSerLeuAlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyr 58
Db 271 ATCGGGGCGCATCTCGGCCATTGTGGCTGTCTCTCTCAGG----- 312
QY 59 TyrPheLeuCysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlu 78
Db 312 ----- 312
QY 79 LeuAspCysProLeuGlyGlnAspGluHisCysValLysSerPheProGluGlyPro 98
Db 313 -----AGTACCAGGAG-----CCG 327
QY 99 AlaValAlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThr 118
Db 328 CTGTACCCAGTGCAGTCTGCGGACGCTCGGCTCATGCTCTTTGACAAGCGGAA 387
QY 119 GlyAsnTrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCys 138
Db 388 GGGACGTGGCGCTGCTGTGCTCTCGCGCTCAACGCCAGGGTAGCGGACTCAGCTGC 447
QY 139 ArgGlnMetGlyTyrSerSerLysProThrPheArgAlaValGluLeuGlyProAspGln 158
Db 448 GAGGAGATGGCTTCCTCAGT----- 468
QY 159 AspLeuAspValValGluLeuThrGluAsnSerGlnGluLeuArgMetArgAsnSerSer 178
Db 468 ----- 468
QY 179 GlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly---LysSer 197
Db 469 ---GATTGCCAGAGCGCGTTCTTGGCGCGCATCTGCAAGACTGTGGCCCGCAGGAAG 525
QY 198 LeuLysThrProArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGln 217


```
Db 526 CTGCGCGTGACCGCATCTGTTGGAGGCGGGACACACAGCTTGGGCGCGTGGCGTGGCAA 585
Qy 218 ValSerIleGlnTyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrp 237
Db 586 GTGAGCTTCGCTATGATGAGACACACCTCTGTGGGGATCCCTGTCTCCGGGAGACTGG 645
Qy 238 ValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPheAsn---TrpLysVal 256
Db 646 GTGCTGACAGCGCCCACTGCTTCCGGAGCGGACCGGGTCTCTGCCGATGGCGAGTG 705
Qy 257 ArgAlaGlySerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIle 276
Db 706 TTGCGCGTGCCTGCGCCAGGCGCTCTCCCAAGTCTGCACTGGGGTGCAGGCTGTG 765
Qy 277 IleGluPheAsnProMetTyrPro-----LysAspAsnAsp 288
Db 766 GTCTACACCGGGGCTATCTTCCCTTCGGGACCCCAACAGCGAGGAGAACAGCAACGAT 825
Qy 289 IleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArgProIleCys 308
Db 826 ATGGCCCTGGTCCACCTCCAGTCCCTGCGCCCTCACAGAATACATCCAGCCTGTGTGC 885
Qy 309 LeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGly 328
Db 886 CTCCAGCTGCGCGGCGCCCTGGTGGATGGCAAGATCTGTACCGTACGCGGCTGGGCGC 945
Qy 329 PheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnVal 348
Db 946 AACACG---CAGTACTATGGCCAAACAGCGCGGGTACTCCAGGAGCTCGAGTCCCAT 1002
Qy 349 IleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyValThrGluLysMet 368
Db 1003 ATCAACATGATGCTCGAATGGCGCTGACTTCTATGGAACCAAGATCAAGCCCAAGATG 1062
Qy 369 MetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyPro 388
Db 1063 TTCTGTGCTGCTACCCCGAGGGTGGCATTGATGCTGCCAGGGCGAGCGGGTGGTCCC 1122
Qy 389 LeuMetTyrGln-----SerAspGlnTrpHisValValGlyValSer 403
Db 1123 TTTGTGTGTAGGACAGCATCTCTCGAGCGCCAGCTTGGCGGCTGTGTGGCATTGTGAGT 1182
Qy 404 TrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyr 423
Db 1183 TGGGCACTGGCTGTGCCCTGGCCCAAGCCAGCGCGTCTACACCAAGTCAAGTCACTTC 1242
Qy 424 LeuAsnTrpIleTyrAsnValTrpLys 432
Db 1243 CGGGAGTGGATCTTCCAGGCCATAAG 1269
```

RESULT 4

```
US-11-067-811-3
; Sequence 3, Application US/11067811
; Publication No. US2005026068A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; PRIOR FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 4804
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(3339)
```

US-11-067-811-3

```
Alignment Scores:
Pred. No.: 6,92e-43 Length: 4804
Score: 488.50 Matches: 130
Percent Similarity: 46.55% Conservative: 52
Best Local Similarity: 33.25% Mismatches: 148
Query Match: 20.86% Indels: 61
DB: 7 Gaps: 17
```

US-10-803-530-2 (1-435) x US-11-067-811-3 (1-4804)

```
Qy 69 IleProArgLysGlnLeuCysAspGlyGluLeuAspCysProLeuGlyGluAspGluGlu 88
Db 2206 GTGCGCGGTGACCTTTGGTGGCAGGATGGTGCAGCTGCTCAGACAGTCTCTGATGAATGG 2265
Qy 89 HisCysValLysSerPheProGluGlyProAlaValAlaValArgLeuSerLysAspArg 108
Db 2266 GGCTGTGTGACCCCTCTCTAAATAATGG-----AACTCTCTCC 2301
Qy 109 SerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrpPheSerAlaCysPheAspAsn 128
Db 2302 TCATTGTGACTGTTCAAAATCTGCAAGGAACAC-----CACGTGTGTGCTGACGGC 2355
Qy 129 PheThrGluAlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSerLysProThr 148
Db 2356 TGGCGGAGAGCTTGAGTCHGCTGGCTGCAAGCAGATGGTTTAGGAGAACCTCTCTGTG 2415
Qy 149 PheArgAlaValGluIleGlyProAspGlnAspLeuAspValValGluIleThrGluAsn 168
Db 2416 ACCAAGCTGATC-----CCAGCAGCAG-----GAAGGC 2442
Qy 169 SerGlnGluLeuArgMet-----ArgAsnSerSerGly----- 179
Db 2443 CAGCAGTGGCTGAGGTGTATCCCAACTCGGAGAAATCTCAATGGGAGCACCTTGACAGGAG 2502
Qy 180 -----ProCysLeuSerGlySerLeuValSerLeuHisCysLeuAla 193
Db 2503 CTGCTGGTATACAGGCACCTCTGCCCAAGCAGAGTGAATTTCCCTTCTGTGCTCCAAG 2562
Qy 194 -----CysGlyLys-----SerLeuLysThrProArgValValGlyGlyGluGlu 208
Db 2563 CAAGACTGTGGCGCGCGCCCTGCTGCCGAATGAACAGAGGATCCTTGGGGTTCGGACT 2622
Qy 209 AlaSerValAspSerTrpProTrpGlnValSerIleGlnTyrAspLysGln---HisVal 227
Db 2623 AGTCGCTCGGAGGTGGCGCGTGGCAGTCTCTCTGCAGAGTGAACCCAGTCGACATATC 2682
Qy 228 CysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLys 247
Db 2683 TGTGGCTGTCTCTCATTTGCCAAGAGTGGTCTCTGACAGTTGCCATTGCTTTGAAGGG 2742
Qy 248 HisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeu-----Gly 264
Db 2743 AGAGAAGACGCTGATGTTTGGAAAGTGTATTTTGGCATAAACAACCTGGACCATCATCA 2802
Qy 265 SerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPro 284
Db 2803 GGCTTCATGCAGACCGCGCTTGTGAAGACCATCTCTGCTA-----CATCCCCGTACAGT 2856
Qy 285 Lys-----AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 2857 CGAGCAGTGGTAGACTATCATATATCAGCGTGTGGAGCTGAGCGATGATATCAATGAGACA 2916
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGluLeuThrProAlaThrPro 321
Db 2917 AGCTACGTACAGCTGTCTGCTACCCAGTCCGAGGAGTATCTAGAACCAACGATACGTAC 2976
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 2977 TGCTACATCACAGGCTGGGCG-----CACATGGCCATAAAATGCCCTTTAAGCTG 3027
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyr--- 360
```

```

Db 3028 CAGGAGGAGAGGTCGCGCATTTATCCCTCTGAGCAGTGC-----CAGTCTCTATTTT 3078
Qy 361 ---GInGlyGluValThrGluLysMetCysAlaGlyIleProGluGlyGlyValAsp 379
Db 3079 GACATGAAGACCATCACCAATCGGATGATCTGTCTGCTATGATCTGGCAGCGTGGAC 3138
Qy 380 ThrCysGlnGlyAspSerGlyProLeuMetTyrGln-----SerAspGlnTrpHis 397
Db 3139 TCCTGCATGGGAGACAGCGGTGGGCTCTGCTGTGTGAACACCGGAGGACAGTGGACA 3198
Qy 398 ValValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThr---ProGlyVal 416
Db 3199 TTATTTGGTTAACTTCATGGGGCTCCCTGCTCTTTCCAAAGTTCTGGGACCTGGAGTG 3258
Qy 417 TyrThrLysValSerAlaTyrLeuAsnTrpIle 427
Db 3259 TACAGCAATGTCTTACTTTTGTGGCTGGATT 3291

RESULT 5
US-11-102-240-105
; Sequence 105, Application US/111102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Grimaldi,Christopher J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Wood,William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 105
; LENGTH: 2103
; TYPE: DNA
; ORGANISM: Homo Sapien
US-11-102-240-105

Alignment Scores:
Pred. No.: 6,42e-42 Length: 2103
Score: 475.00 Matches: 97
Percent Similarity: 56.33% Conservative: 41
Best Local Similarity: 39.58% Mismatches: 97
Query Match: 20.28% Indels: 10
DB: 7 Gaps: 5

US-10-803-530-2 (1-435) x US-11-102-240-105 (1-2103)

Qy 190 HisCysLeuAlaCysGly-----LysSerLeu---LysThrProArgValVal 204
Db 557 CATTGC-----TGGCGAACACGAGAGTAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTT 610
Qy 205 GlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTyrAspLys 224
Db 611 GGTGGGACAGAGTAGAGAGGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGATGGG 670
Qy 225 GlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCys 244
Db 671 AGTCATCGCTGTGGAGCAACCTTAATTAATGACCATGGCTTGTGAGTGCTGCTCACTGT 730
Qy 245 PheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGly 264

```

```

Db 731 TTTACAACATATAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACAATAAAACCT 790
Qy 265 SerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnProMetTyrPro 284
Db 791 TCGAAAATGAACGGGTCTCCGGAGAATAATTTCCATGAAATAACAAACACCCATCA 850
Qy 285 LysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVal 304
Db 851 CATCACTATGATATTTCTCTTGACAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTA 910
Qy 305 ArgProIleCysLeuProPhePheAspGluLeuThrProAlaThrProLeuTrpIle 324
Db 911 CATAGAGTTTCTCCCTCATGATCATCTATGATTTCAACAGGTGATGTGATGTTGTG 970
Qy 325 IleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAla 344
Db 971 ACAGGATTTGGAGCACTGAAA--AATGATGTTACAGTCAAAATCATCTTCGACAAGCA 1027
Qy 345 SerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluVal 364
Db 1028 CAGGTGACTCTCATAGACGCTACAACTTGCAATGAACCTCAAGCTTACAATGACGCCATA 1087
Qy 365 ThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 1088 ACTCTAGAATGTTATGTGCTGCTCTTACAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG 1147
Qy 385 SerGlyGlyProLeuMetTyrGlnSer-----AspGlnTrpHisValValGlyIleVal 402
Db 1148 TCTGGAGACCACTGGTTAGTTAGTTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG 1207
Qy 403 SerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAla 422
Db 1208 AGCTGGGAGAGATGAATGTGCGAAACCCCAACGCTGGTGTATATCTAGAGTTACGGCC 1267
Qy 423 TyrLeuAsnTrpIle 427
Db 1268 TTGGGGAGCTGGATT 1282

RESULT 6
US-11-137-465-23
; Sequence 23, Application US/11137465
; Publication No. US2005025558A1
; GENERAL INFORMATION:
; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdoch, Paul R.
; APPLICANT: Rizvi, Safia, K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kabnick, Karen
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GP50018
; CURRENT APPLICATION NUMBER: US/11/137,465
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US/10/239,663
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/US01/09226
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 60/192,158
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,668
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/200,166
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-137-465-23

Alignment Scores:
Pred. No.: 2.79e-42 Length: 918

```

Score: 473.50 Matches: 95
 Percent Similarity: 54.98% Conservative: 43
 Best Local Similarity: 37.85% Mismatches: 100
 Query Match: 20.22% Indels: 13
 DB: 7 Gaps: 4

US-10-803-530-2 (1-435) x US-11-137-465-23 (1-918)

Qy 193 AlaCysGlyLysSerLeuLysThrProArgValValGlyGluGluAlaSerValAsp 212
 Db 94 GCCTGGCGCCACCGGAAATTCACGCGTGTGGCGGGGAGTGGAGTGGCGGGGG 153
 Qy 213 SerTrpProTrpGlnValSerIleGlnTrpAspLysGlnHisValCysGlySerIle 232
 Db 154 CGCTGGCCATGGCAGGCGAGCTCGCTGGCTGAGGAGCGCCACCGATGGAGGGAGCTG 213
 Qy 233 LeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAspValPhe 252
 Db 214 CTCAGCGCGCTGGTGGCTCTGGCTGGCGACTGCTTCCAAAGACACTATCCCTCC 273
 Qy 253 AsnTrpLysValArgAlaGlySer-----AspLysLeuGlySer 265
 Db 274 GAGTGGAGCTTCAGCTGGCGAGCTGACTTCCAGGCCAATCTTGGAACTGGCGGCC 333
 Qy 266 PheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnProMetTyrProLys 285
 Db 334 TACAGCAGCTGTTACAAAGTGCAGGACATATTGTAACCTGACCGCACTTGGGGTTTA 393
 Qy 286 AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
 Db 394 CGCAATGACATTGCTGCTGAGACTGGCTCTTCTGTCACCTAATGCGTACATCCAG 453
 Qy 306 ProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleIle 325
 Db 454 CCCATTGTCATGAGTCTTCCACCTTCACTTGTGTCAGCGCGGAGTGTGGTGACC 513
 Qy 326 GlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAsp-----IleLeuLeuGln 343
 Db 514 GGCTGGGGTAAATCAGCGCCAGTGGCACACTCTGCCACCTCTTACAACCTCCGGGA 573
 Qy 344 AlaSerValGlnValIleAspSerThrArgCysAsn-----AlaAspAspAlaTyr 360
 Db 574 GCACAGTGCACACTTAAACACACACAGGCTGTAATTACCTGTTTGAACAGCCCTTAGC 633
 Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
 Db 634 CGTAGTAGTCTGGGATTCATGTTTGTGCTGCTGAGGATGGCAGTGTAGACACC 693
 Qy 381 CysGlnGlyAspSerGlyProLeuMetTyrGlnSerAsp---GlnTrpHisValVal 399
 Db 694 TGCAAAGTGACTCAGGTGACCCCTTGTCTGTGCAAGGATGAGTGTGTATCAGGTT 753
 Qy 400 GlyIleValSerTrpGlyTyrGlyCysGlyProSerThrProGlyValTyrThrLys 419
 Db 754 GGAATCTGTAGCTGGGAATGAGTGGCGGTCAACCCCAATGGCGCTGTGTATCACCAAC 813
 Qy 420 ValSerAlaTyrLeuAsnTrpIleTyrAsnVal 430
 Db 814 ATCAGTGTGACTTCCACTGGATCGGAGGTG 846

RESULT 7

US-10-131-826A-221
 ; Sequence 221, Application US/10131826A
 ; Publication No. US20050245730A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: DeForge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Geurtsen, Mary E.
 ; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C128
 ; CURRENT APPLICATION NUMBER: US/10/131,826A
 ; CURRENT FILING DATE: 2002-04-24
 ; PRIOR APPLICATION NUMBER: 60/049911
 ; PRIOR FILING DATE: 1997-06-18
 ; PRIOR APPLICATION NUMBER: 60/056974
 ; PRIOR FILING DATE: 1997-08-26
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059115
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059117
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059122
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059184
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/059263
 ; PRIOR FILING DATE: 1997-09-18
 ; PRIOR APPLICATION NUMBER: 60/059352
 ; PRIOR FILING DATE: 1997-09-19
 ; PRIOR APPLICATION NUMBER: 60/059588
 ; PRIOR FILING DATE: 1997-09-19
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 221
 ; LENGTH: 1129
 ; TYPE: DNA
 ; ORGANISM: Homo Sapien
 ; US-10-131-826A-221

Alignment Scores:
 Pred. No.: 1,74e-41 Length: 1129
 Score: 467.50 Matches: 99
 Percent Similarity: 55.43% Conservative: 49
 Best Local Similarity: 37.08% Mismatches: 92
 Query Match: 19.96% Indels: 27
 DB: 6 Gaps: 9

US-10-803-530-2 (1-435) x US-10-131-826A-221 (1-1129)

Qy 184 GlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrProArgVal 203
 Db 81 GGCTCTCAGAGGCCAAGGCGAGCAACAGCCTGTGTGCCCCCAGGATGCTGAACCGAATG 140
 Qy 204 ValGlyGlyGluGluAlaSerValAspSerTrpTrpGlnValSerIleGlnTrpAsp 223
 Db 141 GTGGCGGGCAGGACACGCGAGGAGGGAGTGGCCCTGGCAAGTCAGCATCCAGCGCAAC 200
 Qy 224 LysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHis 243
 Db 201 GGAAGCCACTTCTGCGGGGCGACCTCATCGCGAGGAGTGGTCTCTGACGGTGGCGAC 260
 Qy 244 CysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeu 263
 Db 261 TGCTTCGCAACACCTCTGAGACGTCCTGACAGTCTGTGAGGGGCAAGCACTA 320
 Qy 264 ---GlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnProMet 282
 Db 321 GTCCAGCGGGACACACCGCTATGTATGCCGGGTGAGGAGGAGGAGCAACCCCTG 380
 Qy 283 Tyr-----ProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThr 299

```

Db 381 TACCAGGCGACGGGCTCCAGCGCTGACGTGGCCCTGGTGGAGCTGGAGCGCACCGAGTGCCC 440
Qy 300 PheSerGlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAla 319
Db 441 TTCACCAATTACATCTCCCGTGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 494
Qy 320 ThrProLeu-----TrrPheGlyThrPheThrLysGlnAsnGlyGlyLysMet 337
Db 495 ACGGCGATGAATCTGCTGGGTCACTGGCTGGGCGAGCCCGAGTGAG----- 539
Qy 338 SerAsp-----IleLeuLeuGlnAlaSerValGlnValIleAspSer 351
Db 540 GAAGACCTCTCTGCGCGAACCCTGCTGAGAACTCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
Qy 352 ThrArgCysAsn-----AlaAspAspAlaTyrGlnGlyGlu---Val 364
Db 600 CCCAGTGCACCTGCTCTACAGCAAGACACCGAGTTGGCTACCAACCAACCAATC 659
Qy 365 ThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAsp 384
Db 660 AAGATGACATGCTGTGCGCGCTTCGAGGAGGCGCAAGAGGATGCTGCAAGGCGAC 719
Qy 385 SerGlyGlyProLeuMetTyrGlnSerAspGln---TrpHisValValGlyIleValSer 403
Db 720 TCGGCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
Qy 404 TrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyr 423
Db 780 TGGGTGAGGGGTGTGCGCGCGAGAACCGCCAGGTGTCTACATCCGTGTACCGCGCCAC 839
Qy 424 LeuAsnTrpIleTyrAsnVal 430
Db 840 CACAACCTGCATCCATCGGATC 860

```

RESULT 8

```

US-10-821-234-191/c
; Sequence 191, Application US/10821234
; Publication No. US200502551141
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_SEQ_genes Version 1.0
; SEQ ID NO 191
; LENGTH: 1802
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-191

```

Alignment Scores:

```

Pred. No.: 3,87e-39 Length: 1802
Score: 449.00 Matches: 113
Percent Similarity: 49.51% Conservative: 40
Best Local Similarity: 36.57% Mismatches: 116
Query Match: 19.17% Indels: 40
DB: 6 Gaps: 10

```

US-10-803-530-2 (1-435) x US-10-821-234-191 (1-1802)

```

Qy 147 ProThrPheArgAlaValGluIle-----GlyProAspGlnAspLeuAspValValGlu 164
Db 1666 CTACCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1631
Qy 165 IleThrGluAsnSerGlnGluLeuArgMetArgAsnSerGlyProCysLeuSerGly 184

```

```

Db 1630 -----TTGTCTGGGCGCATGGCCAGAAAGGGGGTCTCTGGGCGCTGGGCGAGCTGGGG 1580
Qy 185 SerLeuValSerLeuHisCysLeuAla----- 193
Db 1579 GCTGTGGCCATTCTGCTCTATCTTTGGATTACTCCGGTCGGGACAGGAGCGGAAGGGCA 1520
Qy 194 -----CysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSer 210
Db 1519 GAAGCTCCTCGGTGTGGCCCCCAAGCA---CGCATCACAGGTGGGAGCAGTCAGTC 1463
Qy 211 ValAspSerTrpTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlyGly 230
Db 1462 GCCGTCAGTGGCCCTGGCAGGTGAGCATCACTATGAAGCGCTCCATGTGTGTGTGGTGC 1403
Qy 231 SerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThrAsp 250
Db 1402 TCTCTCGTGTGAGCAGTGGGTGCTGTCAGCTGCTCACTCTCCCGCAGCGAGCACAC 1343
Qy 251 ValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPheProSerLeuAla 270
Db 1342 AAGGAAGCCTATGAGTCAAGCTGGGGGCCCAACAGCTAGACTCTCTACTCCGAGGAGCGCC 1283
Qy 271 ---ValAlaLysIleIleIleGluPheAsnProMetTyrProLysAsp----- 286
Db 1282 AAGTTCAGCACCCCTGAAGGACATCATCCCCCAGCTACCTCCAGGAGGGCTCCCCAG 1223
Qy 287 AsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArgPro 306
Db 1222 GCGGACATGACACTCTCCAACTCAGAGACCCATCACTCTCTCCGCTACATCCGCGCC 1163
Qy 307 IleCysLeuProPhePheAspGluLeuThrProAlaThrProLeuTrpIleIleGly 326
Db 1162 ATCTGCTCTCCCTGCAGCCACGCCCTCTTCCCAAGCGCTCCACTGCACCTGCTCACTGCG 1103
Qy 327 TrpGlyPheThrLysGlnAsnGlyGlyLysMetSer---AspIleLeuLeuGlnAlaSer 345
Db 1102 TGGGTCTATGTGGCCCCCTCAGTGAGCTCTCTGAGCGCCCAAGCCACTGCAGCAACTCGAG 1043
Qy 346 ValGlnValIleAspSerThrArgCysAsnAla-----AspAspAlaTyrGlnGly 362
Db 1042 GTGCTCTGATCAGTCGTGAGACGTGTAACCTGCTACCAATCGAGCCCAAGCCTGAG 983
Qy 363 Glu-----ValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAsp 379
Db 982 GAGCGCGACTTGTCCAGAGGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 923
Qy 380 ThrCysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAsp---GlnTrpHisVal 398
Db 922 GCCTGCCAGGGTGACTCTGGGGGCCCACTCTCTGCTGCTGTGGAGGGTCTCTGTGTACCTG 863
Qy 399 ValGlyIleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThr 418
Db 862 ACGGGCATTTGAGCTGGGAGATGCTGTGGGGGCCGCAACAGGCGCTGTGTGTGTACT 803
Qy 419 LysValSerAlaTyrLeuAsnTrpIle 427
Db 802 CTGGCTCCAGCTATGCTCTCTGGATC 776

```

RESULT 9

```

US-11-056-621-5
; Sequence 5, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gasdaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; TITLE OF INVENTION: MICROPLASMINOGEN IN DUCKWEED
; FILE REFERENCE: 40989/274646

```

```

; CURRENT APPLICATION NUMBER: US/11/056,621
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Duckweed codon optimized sequence encoding human
; OTHER INFORMATION: microplasma
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-11-056-621-5

Alignment Scores:
Pred. No.: 5,936-39 Length: 780
Score: 442.50 Matches: 97
Percent Similarity: 52.61% Conservative: 44
Best Local Similarity: 36.19% Mismatches: 98
Query Match: 18.89% Indels: 29
DB: 7 Gaps: 7

US-10-803-530-2 (1-435) x US-11-056-621-5 (1-780)
QY 176 AsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly 195
DB 22 AACGCGAGGCGCGTCTTCGAC-----ArgValValGlyGlyGluGluAlaSer 210
QY 196 LysSerLeuHisThrPro-----ArgValValGlyGlyGluGluAlaSer 210
DB 52 AAGCCACAGGTGGACCGAAGTGCCTGGCGCGGTGTCGAGGGTGCCTGGCCAC 111
QY 211 ValAspSerTrpProTrpGlnValSerIleGlnTyrAsp---LysGlnHisValCysGly 229
DB 112 CCGCAGCTCTGGCGCTGCAAGTCAGCTCGCGACCGCTTCGGCATGCATCTTCGGCG 171
QY 230 GlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThr 249
DB 172 GGCACCTCATCTCCCGAGTGGTTCAGCGCGCTCCTGCTCGAGAGTCCCGG 231
QY 250 AspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhe----- 266
DB 232 AGGCCCTCTCTACAGGTCTATCTGGCGGCCCGCCAGAGGTGAACCTCGAGCGCGAC 291
QY 267 ---ProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrProLys 285
DB 292 GTTCAGGAGATCAGGTGTCTCCGCTTGTTC-----CTGGAGGCCACG 333
QY 286 AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
DB 334 CGCAAGATATCGCCCTGCTCAAGCTCTTAGCCGCGCGTCACTACCGACAAAGTTATC 393
QY 306 ProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIle 325
DB 394 CCGCGCTCTCTCCCTCCCGAAGTACGTGCTGCGTACCGACCGAGGTCTTCCTTACC 453
QY 326 GlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeuLeuGlnAlaSer 345
DB 454 GGCTGGGGGAGACCCAGGAAAGTTCGCG-----GCGGGCTCTCAAGGAGGCCACG 507
QY 346 ValGlnValIleAspSerThrArgCysAnAlaAspAlaTyrGlnGlyGluValThr 365
DB 508 CTCCCGGTGATTGAGAACAAAGTGTGCAACCGTTACGAGTTCCTGAACGGCGCGTCCAG 567
QY 366 GluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSer 385
DB 568 TCCACCGAAGTCTGCGCGGGCACTTGGCGCGCGACCGACAGCTGCCAGGGCGACG 627
QY 386 GlyGlyProLeuMet---TyrGlnSerAspGlnTrpHisValValGlyIleValSerTrp 404
; CURRENT APPLICATION NUMBER: US/11/056,621
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 780
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Duckweed codon optimized sequence encoding human
; OTHER INFORMATION: microplasma
; NAME/KEY: CDS
; LOCATION: (1)...(780)
US-11-056-621-5

Alignment Scores:
Pred. No.: 5,936-39 Length: 780
Score: 442.50 Matches: 97
Percent Similarity: 52.61% Conservative: 44
Best Local Similarity: 36.19% Mismatches: 98
Query Match: 18.89% Indels: 29
DB: 7 Gaps: 7

US-10-803-530-2 (1-435) x US-11-056-621-5 (1-780)
QY 176 AsnSerSerGlyProCysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGly 195
DB 22 AACGCGAGGCGCGTCTTCGAC-----ArgValValGlyGlyGluGluAlaSer 210
QY 196 LysSerLeuHisThrPro-----ArgValValGlyGlyGluGluAlaSer 210
DB 52 AAGCCACAGGTGGACCGAAGTGCCTGGCGCGGTGTCGAGGGTGCCTGGCCAC 111
QY 211 ValAspSerTrpProTrpGlnValSerIleGlnTyrAsp---LysGlnHisValCysGly 229
DB 112 CCGCAGCTCTGGCGCTGCAAGTCAGCTCGCGACCGCTTCGGCATGCATCTTCGGCG 171
QY 230 GlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisThr 249
DB 172 GGCACCTCATCTCCCGAGTGGTTCAGCGCGCTCCTGCTCGAGAGTCCCGG 231
QY 250 AspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhe----- 266
DB 232 AGGCCCTCTCTACAGGTCTATCTGGCGGCCCGCCAGAGGTGAACCTCGAGCGCGAC 291
QY 267 ---ProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrProLys 285
DB 292 GTTCAGGAGATCAGGTGTCTCCGCTTGTTC-----CTGGAGGCCACG 333
QY 286 AspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrValArg 305
DB 334 CGCAAGATATCGCCCTGCTCAAGCTCTTAGCCGCGCGTCACTACCGACAAAGTTATC 393
QY 306 ProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIle 325
DB 394 CCGCGCTCTCTCCCTCCCGAAGTACGTGCTGCGTACCGACCGAGGTCTTCCTTACC 453
QY 326 GlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeuLeuGlnAlaSer 345
DB 454 GGCTGGGGGAGACCCAGGAAAGTTCGCG-----GCGGGCTCTCAAGGAGGCCACG 507
QY 346 ValGlnValIleAspSerThrArgCysAnAlaAspAlaTyrGlnGlyGluValThr 365
DB 508 CTCCCGGTGATTGAGAACAAAGTGTGCAACCGTTACGAGTTCCTGAACGGCGCGTCCAG 567
QY 366 GluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSer 385
DB 568 TCCACCGAAGTCTGCGCGGGCACTTGGCGCGCGACCGACAGCTGCCAGGGCGACG 627
QY 386 GlyGlyProLeuMet---TyrGlnSerAspGlnTrpHisValValGlyIleValSerTrp 404

```

```
QY 195 lyLysSerLeuLysThrPro-----ArgValValGlyGlyGluGluAla 210
Db 1646 GGAAGCCACAGTGGAAACCGAAGAAGTGCCTGGCGCGTGTGTGGAGGGTGGTGGCC 1705
QY 210 erValAspSerTrpProTrpGlnValSerIleGlnTyrAsp---LysGlnHisValCysG 229
Db 1706 ACCCGCACTCTGGCGCCCTGGCAAGTGCAGTCCGCCAGCCCGTTCGGCATGCACTTCTCGG 1765
QY 229 lyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLysHisIst 249
Db 1766 GCGGACCTCATCTCCCGGAGTGGTTCAGCCGCGCTCACTGCTCGAGAGTCC 1825
QY 249 hrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySerPhe----- 266
Db 1826 CGAGCCCTCTCTACAGGTGCATCTGGGGGCCACCGAGGAGTGAACCTCGAGCCGC 1885
QY 267 -----ProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrProL 285
Db 1886 AGTTTCAGGAGATCGAGGTGCTCCCGTTGTC-----CTGGAGCCCA 1927
QY 285 ysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSerGlyThrVala 305
Db 1928 CCGCAAGATATCCCGCTCTCAAGCTCTTAGCCCGCGCGTCTATCACCGAAGGTGA 1987
QY 305 rgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrProLeuTrpIleI 325
Db 1988 TCCCGGCTGCTTCCCTCCCGAAGTGTGCTGCTGACCGCACCGAGTCTTCGTGA 2047
QY 325 leGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlas 345
Db 2048 CCGGCTGGCGGAGACCCAGGAGACGTTCTGGC-----GCGGCGCTCTCAAGGAGGCC 2101
QY 345 erValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyGluValT 365
Db 2102 AGCTCCCGGTATTGAGAACAGAGGTGTGCACCGTTACAGTTCCTCAACGCGCGCTCC 2161
QY 365 hrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspS 385
Db 2162 AGTCACCGAACTCTGCGCGCGGCACTTGGCGCGGCGCACCGACAGCTGCCAGGCGGACA 2221
QY 385 erGlyGlyProLeuMet---TyrGlnSerAspGlnTrpHisValValGlyIleValSerT 404
Db 2222 GCGCGGCGCGCTGTGTGTCTTCGAGAGGACAGTACATCTCCAGCGCGTCAGCTCT 2281
QY 404 rgGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrL 424
Db 2282 GGGGCTCGGCTGGCAGCCCTTAACAGCGGGGCTATGTGCGCGGTCTCCCGTTCG 2341
QY 424 euAsnTrpIleTyrAsnValTrpLys 432
Db 2342 TGACCTGGATCGAGGCGGTGATGCGC 2367
```

RESULT 11

```
US-10-131-826A-505
; Sequence 505, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 505
; LENGTH: 1204
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-505

Alignment Scores:
Pred. No.: 1,76e-31 Length: 1204
Score: 377.50 Matches: 99
Percent Similarity: 49.15% Conservative: 45
Best Local Similarity: 33.79% Mismatches: 111
Query Match: 16.12% Indels: 38
DB: 6 Gaps: 11

US-10-803-530-2 (1-435) x US-10-131-826A-505 (1-1204)
QY 163 ValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSer----- 178
Db 18 GTTAGGTGGCTGCGGGACTGGAAGTCTATCGGGCAGAGGTCTCACAGCAGCCAAGGAACC 77
QY 179 ---GlyProCysLeu-----SerGlySerLeuValSerLeuHisCys-Le 192
Db 78 TGGGGCGCGCTCTCTCCCGCTCCAGGCCATGAGGATTCTGCAGTTAATCTGCTGCTCT 137
QY 192 uAlaCysGlyLysSerLeuLysThrProArgValValGlyGlyGluGluAlaSerVala 212
Db 138 GGCACACAGGCTTGTAGGGGGAGAGACCAGGATATCATCAAGGGGTTCGAGTGCAGCCTCA 197
QY 212 pSerTrpProTrpGlnValSerIleGlnTyrAspLysGlnHisValCysGlySerI 232
Db 198 CTCCAGCCCTGGCAGGCGCCCTGTTTCGAGAGAGCGCGCTACTCTGTGGGCGGCGCT 257
QY 232 eLeuAspProHisTrpValLeuThrAlaAlaHisCysPheArgLys-----Hi 248
Db 258 CATCCCGCCAGATGGCTCTCTGACAGCAGCCCACTGCTCAAGCCCGCTACATAGTTCA 317
QY 248 sThrAspValPheAsnTrpLysValArgAlaGlySerAspLys-----LeuGl 264
Db 318 CTGGGGCAGCACAACTCCAGAGGAGGAGGCGCTGTGAGCAGAGCCCGCAGCAGCCACTGA 377
QY 264 ySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnProMetTyrPr 284
Db 378 GTCTCTCCCGCCCGCGGC-----TTCAACAACAGCCTCCC 413
QY 284 o---LysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSe 301
```

```
Db 414 CAACAAAGACCCAGCAATGACATCATGTGTGAAGATGGCATCGCCAGTCTCCATCAC 473
QY 301 rGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPr 321
Db 474 CTGGGCTGTGGCAGCCCTCACCTCTCC-----TCACGCTGTGTCACTGTGGCACCAG 527
QY 321 oLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLe 341
Db 528 CTGCTCATTTCCGGCTGGGGCAGCAGCTGCAGCCCCCAGTTACCCCTGCCTCACACCTT 587
QY 341 uLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrG1 361
Db 598 GCGATGCGCCACATCACCATCATTTGAGCAGCAGAGTGT-----GAGAAGCGCTACCC 641
QY 361 nGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCy 381
Db 642 CGCGAACATCACAGACCATGTGTGTGCCAGCGTGCAGGAAGGGGCAAGGACTCTCTG 701
QY 381 sGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyI1 401
Db 702 CCAGGCTGACTCCGGGGCCCTCTGTGTCTGTAAACAGTCT-----CTTCAAGGCAT 752
QY 401 eValSerTrpGlyTyrGly---CysGlyGlyProSerThrProGlyValTyrThrLysVa 420
Db 753 TATCTCTGGGCGCAGGATCCGTGTGGCATCACCCGAAAGCGTGGTGTCTACACGAAAGT 812
QY 420 lSerIleTyrLeuAsnTrpIleTyrAsnValTrpIys 432
Db 813 CTGCAAAATATGTGGACTGGATCCAGGAGCATGAAG 849
```

RESULT 12

```
US-11-029-003-7
; Sequence 7, Application US/11029003
; Publication No. US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.
; APPLICANT: BITONTI, ALAN J.
; APPLICANT: STATTLE, JAMES
; TITLE OF INVENTION: IMMUNOGLOBULIN CHIMERIC MONOMER-DIMER HYBRIDS
; FILE REFERENCE: 08945.0007-01000
; CURRENT APPLICATION NUMBER: US/11/029,003
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 60/539,207
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: 60/487,964
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/469,600
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 7
; LENGTH: 2016
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-11-029-003-7
```

```
Alignment Scores:
Pred. No.: 2,88e-30 Length: 2016
Score: 369.50 Matches: 118
Percent Similarity: 41.65% Conservative: 54
Best Local Similarity: 28.57% Mismatches: 127
Query Match: 15.78% Indels: 114
DB: 7 Gaps: 16
```

US-10-803-530-2 (1-435) x US-11-029-003-7 (1-2016)

QY 84 GlyGluAspGlu-----GluHisCysValLysSer 93

```
Db 216 GGAGAGACCAAGCTGTCTGGATTCTTACAGTATGGGACCAAGTGTGCTCAAGTCC 275
QY 94 PheProGlu-----GlyProAlaValAlaValArgLeu----- 104
Db 276 ATGCCAGAATGGGGCTCTTCAAGGACCAAGCTCCAGTCTATATCTGCTTCTGCTCC 335
QY 105 -----SerLys 106
Db 336 TGCTTCGAGGGCGGAACTGTGAGACGCACAAGGATGACCAAGCTGATCTGTGTGAACGA 395
QY 107 AspArgSerThrLeuGlnValLeu-AspSerAlaThrGlyAsnTrpPheSerAla---Cy 125
Db 396 GAACGGCGGTGTGAGCAGTACTGACAGTACCAACACGGGACCAAGCGCTCTGTCTCGGTG 455
QY 125 sPheAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMetGlyTyrSerSe 145
Db 456 CCACGAGGGTACTCT---CTGCTGGCAGAC-----GGGGTGTCTCTG 494
QY 145 rLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAspValValGluI1 165
Db 495 CACACCCACA-----GTTGAATAT---CCATGTGGAAAAAATACCTATTCTAGAA-- 540
QY 165 eThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCysLeuSerGlySe 185
Db 541 -----AAAGAAATGCCAGCAAAACCC----- 561
QY 185 rLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrProArgValValG1 205
Db 562 -----CAAGGCGGAATTGTGGG 578
QY 205 YGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGlnTyrAspLysG1 225
Db 579 GGGCAAGGTGTGCCCAAGGGGAGTGTCCATGCGAGGTCTCTGTGTGTGTAATGGAGC 638
QY 225 nHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPh 245
Db 639 TCAGTTGTGTGGGGGAGCCCTGATCAACACCATCTGGTGTCTCGCGCGCCACTGTTT 698
QY 245 eArgLysHisThrAspValPheAsnTrpLysValArgAlaGly----- 259
Db 699 CGACAAA-----ATCAAGAACTGGAGAACTGATCGCGGTCTGGCGGACGACGA 749
QY 260 ----SerAspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleG1 278
Db 750 CCTCAGCGAGCACGACGGGATGAGCAGAGCCGCGGTGGCGAGTCAATCATCCCCAG 809
QY 278 uPheAsnProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLe 298
Db 810 CACGTACGTCCCGGCGCACCAACCAACACGACATCGCGCTGCTCCGCTGCACACGCGCGT 869
QY 298 uThrPheSerGlyThrValArgProIleCysLeuPro-----PhePheAspGluG1 315
Db 870 GGTCTCCTACCTACCATGTGTGTCCTCTGCTCCCGAAGCAGGAGCTTCTCTGAGGAGAC 929
QY 315 uLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyG1 335
Db 930 GCTGGCTTCTGTCGCTTCTCATTTGTCAGCGGCTGGGGCCAGCTGCTGGACCGCTGGCGC 999
QY 335 YLysMetSerAspIleLeu-----LeuG1 343
Db 990 CACGGCCCTGGAGCTCATGGTCTCTCAACGTGCCCGGCTGATGACCCAGGAGCTGCCTGCA 1049
QY 343 nAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGlnGlyG1 363
Db 1050 GCAGTCACGAAGTGGGAGACTCCCA-----AA 1079
QY 363 uValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnG1 383
Db 1080 TATCAGGAGTACATGTTCTGTGCGCGCTACTCGGATGGCAGCAAGGACTCTCTGCAAGGG 1139
QY 383 YAspSerGlyGlyPro---LeuMetTyrGlnSerAspGlnTrpHisValValGlyIleVa 402
```



```
Db 1140 GGACAGTGGAGGCCACATGCCACCCACTACCGGGGCAGTGGTACCTGACGGGCATCGT 1199
Qy 402 lserTrpGlyTyGlyCysGlyProSerThrProGlyValTyThrLysValSerAl 422
Db 1200 CAGCTGGGGCCAGGCGTGGCAACCGTGGGCCACTTTTGGGGTGTACACCGAGGCTCCCA 1259
Qy 422 atyLeuAenTrrpIleTyAenValTrpLysAlaGlu 434
Db 1260 GTACATCGAGTGGCTGCCAAAGACTCATGGCTCAGAG 1296

RESULT 13
US-10-623-155-122
; Sequence 122, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C20
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 1475
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-122

Alignment Scores:
Pred. No.: 5,686-29 Length: 1475
Score: 356.00 Matches: 91
Percent Similarity: 49.25% Conservative: 40
Best Local Similarity: 34.21% Mismatches: 103
Query Match: 15.20% Indels: 32
DB: Gaps: 9

US-10-803-530-2 (1-435) x US-10-623-155-122 (1-1475)
Qy 194 CysGlyLysSerLeuLysThrProArg-----ValValGlyGlyGluAlaSerVal 211
Db 582 TGTGCCAAAGAGACTCTGAGGCCCGCTTTAAGATTATTGGGGGAGAAATTCACCAACATC 641
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyAspLysGln----- 225
Db 642 GAGAACCCAGCCCTGTTTGGCGCCATC---TACAGGAGGACCGGGGGGCTGTGTACC 698
Qy 226 HisValCysGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe 245
Db 699 TAGGTGTGTGAGGAGCGCTCATCAGCCCTTCTGGGTGATCAGCGCCACACACTGCTTC 758
Qy 246 ArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySer 265
Db 759 ATTGATTACCCAAAGAGGAGACTACATCGTCTACCTGGTGGTGGTCAAGGCTTAACCTCC 818
Qy 266 PhePro-----SerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro 281
Db 819 AACACGCAAGGGGAGATGAAGTTTGAGGTGGAAACCTCATCTCTA-----CACAG 869
Qy 282 MetTyProLysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPhe 296
Db 870 GACTACAGCGCTGACACGCTTGTCTACCAACAGCATTTGCTTCTGAAGATCCGTTCC 929
Qy 297 -----ProLeuThrPheSerGlyThrValArgProIleCysLeu 309
Db 930 AAGGAGGCGAGGTGTGCGCAGCCA-----TCCCGACTATACAGACCATCTGCTG 980
Qy 310 ProphePheAspGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPhe 329
Db 981 CCCTCGATGATATACGATCCCGAGTTTGGCACAAAGCTGTGAGATCATCTGGCTTTGAAAA 1040
```

```
Qy 330 ThrLysGlnAenGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIle 349
Db 1041 GAGAATTCACCGACTATCTCTATCCGAGCAGCTGAAAGATGACTGTTGTGAAGCTGATT 1100
Qy 350 AspSerThrArgCysAsnAlaAspAlaTyTrpGlnGlyGluValThrGluLysMetMet 369
Db 1101 TCCACACCGGAGTGTACGACGCCCTACTACGGCTCTGAAGTCCACCAACCAATGCTG 1160
Qy 370 CysAlaGlyIleProGluGlyValAspThrCysGlnGlyAspSerGlyGlyProLeu 389
Db 1161 TGTGCTGTGACCCACAGTGGAAAAACAGATTCTCCAGGGAGACTCAGGGGGACCCCTC 1220
Qy 390 MetTyGln---SerAspGlnTrpHisValValGlyIleValSerTrpGlyTyGlyCys 408
Db 1221 GTCTGTTCCCTCCAAAGCCGCATGACTTTGAGTGAATTTGAGTGGGGCGGTGGATGT 1280
Qy 409 GlyGlyProSerThrProGlyValTyThrLysValSerAlaTyTrpLeuAsnTrpIleTy 428
Db 1281 GCCCTGAAGGACAGCCAGGCGTCTACACGAGAGTCTCACACTTCTTACCTTGGATCCGC 1340
Qy 429 AsnValTrpLysAlaGlu 434
Db 1341 AGTCACACCAAGGAAGAG 1358

RESULT 14
US-10-623-155-123
; Sequence 123, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Peckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C20
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 123
; LENGTH: 2294
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-623-155-123

Alignment Scores:
Pred. No.: 1,098-28 Length: 2294
Score: 356.00 Matches: 91
Percent Similarity: 49.25% Conservative: 40
Best Local Similarity: 34.21% Mismatches: 103
Query Match: 15.20% Indels: 32
DB: Gaps: 9

US-10-803-530-2 (1-435) x US-10-623-155-123 (1-2294)
Qy 194 CysGlyLysSerLeuLysThrProArg-----ValValGlyGlyGluAlaSerVal 211
Db 582 TGTGCCAAAGAGACTCTGAGGCCCGCTTTAAGATTATTGGGGGAGAAATTCACCAACATC 627
Qy 212 AspSerTrpProTrpGlnValSerIleGlnTyAspLysGln----- 225
Db 628 GAGAACCCAGCCCTGTTTGGCGCCATC---TACAGGAGGACCGGGGGGCTGTGTACC 684
Qy 226 HisValCysGlySerIleLeuAspProHisTrpValLeuThrAlaAlaHisCysPhe 245
Db 685 TAGGTGTGTGAGGAGCGCTCATCAGCCCTTCTGGGTGATCAGCGCCACACACTGCTTC 744
Qy 246 ArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAspLysLeuGlySer 265
Db 745 ATTGATTACCCAAAGAGGAGACTACATCGTCTACCTGGTGGTGGTCAAGGCTTAACCTCC 804
```

```
QY 266 PhePro-----SerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db      :::::::::::
QY 805 AACACGCAAGGGAGATGAAGTTGAGTGGAAAAACCTAATCCTA-----CACAAG 855
Db      :::::::::::
QY 282 MetTyrProLysAsp-----AsnAspIleAlaLeuMetLysLeuGlnPhe 296
Db      :::::::::::
QY 856 GACTACAGCGGTGACACGCTGCTCACCACACAGCATTCCTTGGTGAAGATCCGTTCC 915
Db      :::::::::::
QY 297 -----ProLeuThrPheSerGlyThrValArgProIleCysLeu 309
Db      :::::::::::
QY 916 AAGGAGGCGAGGTGCGCAGCCA-----TCCCGGACTATACAGACCATCTGCCTG 966
Db      :::::::::::
QY 310 ProPheAspGluGluLeuThrProAlaThrProLeuTrpIleIleGlyTrpGlyPhe 329
Db      :::::::::::
QY 967 CCCTCGATGATATACGATCCCGAGTTGGCACAAGCTGTGAGATCACTGGCTTTGAAAA 1026
Db      :::::::::::
QY 330 ThrLysGlnAsnGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnValIle 349
Db      :::::::::::
QY 1027 GAGAATCTACCGACTATCTATCCGAGCAGCTGAAATGACTGTTGTGAAGCTGAT 1086
Db      :::::::::::
QY 350 AspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMetMet 369
Db      :::::::::::
QY 1087 TCCACCGGAGGTGTACAGACCCCACTACTACGGCTCTGAAGTCAACCAAAATGCTG 1146
Db      :::::::::::
QY 370 CysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyProLeu 389
Db      :::::::::::
QY 1147 TGTGCTGCTGACCCACAGTGGAAACAGATTCCTGCAGGAGACTCAGGGGACCCCTC 1206
Db      :::::::::::
QY 390 MetTyrGln---SerAspGlnTrpHisValValGlyIleValSerTrpGlyCys 408
Db      :::::::::::
QY 1207 GTCTGTTCCTCCCAAGCCCGCATGACTTTGACTGGAATTTGTGAGCTGGGCGGTGAT 1266
Db      :::::::::::
QY 409 GlyGlyProSerThrProGlyValTyrThrLysValSerAlaTyrLeuAsnTrpIleTyr 428
Db      :::::::::::
QY 1267 GCCTGAGACACAGCCGCGTCTACAGAGAGTCTCACACTTCTTACCTGGATCCGC 1326
Db      :::::::::::
QY 429 AsnValTrpLysAlaGlu 434
Db      :::::::::::
QY 1327 AGTCACACCAAGGAAG 1344
Db      :::::::::::
RESULT 15
US-11-029-003-9
; Sequence 9, Application US/11029003
; Publication No. US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.
; APPLICANT: BITONTI, ALAN J.
; APPLICANT: STATTEL, JAMES
; TITLE OF INVENTION: IMMUNOGLOBULIN CHIMERIC MONOMER-DIMER HYBRIDS
; FILE REFERENCE: 08945.0007-01000
; CURRENT APPLICATION NUMBER: US/11/029,003
; PRIOR FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 60/539,207
; PRIOR FILING DATE: 2004-01-26
; PRIOR APPLICATION NUMBER: 60/487,964
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/469,600
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 9
; LENGTH: 2091
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-11-029-003-9
Alignment Scores:
Pred. No.: 1.08e-28 Length: 2091
```

```
Score: 355.50 Matches: 109
Percent Similarity: 41.84% Conservative: 50
Best Local Similarity: 28.68% Mismatches: 144
Query Match: 15.18% Indels: 77
DB: 7 Gaps: 15
US-10-803-530-2 (1-435) x US-11-029-003-9 (1-2091)
QY 81 CysProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaVal 100
Db      :::::::::::
QY 355 TGTCCCTTTGGATTGAGGAAAGAACTGTGAATTAGAT-----GTA 396
Db      :::::::::::
QY 101 AlaValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
Db      :::::::::::
QY 397 ACATGTACATTAAGAATGCGCAGATGCGAGAGTTTGTAAATAATAGTGTGTATAACAAG 456
Db      :::::::::::
QY 121 TrpPheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArg--- 139
Db      :::::::::::
QY 457 GTGGTTTTCCTCTACTAGGATATCGACTTGCAGAAACACCAAGTCTGTGAACCA 516
Db      :::::::::::
QY 140 -----GlnMetGlyTyrSerSerLysProThrPheArg 150
Db      :::::::::::
QY 517 GCAGTGCCATTTCCATGTGGAAGAGTTTCTGTTTTCACAACTTCTAAGCTCACC---CGT 573
Db      :::::::::::
QY 151 AlaValGluIleGlyProAspGlnAsp-----Leu 160
Db      :::::::::::
QY 574 GCTGAGAGCTGTTTCTGATGTGGACTATGTAAATTTCTACTGAAGCTGAACCACTTTTG 633
Db      :::::::::::
QY 161 AspValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyPro 180
Db      :::::::::::
QY 634 GAT-----AACATCACTCAAGCACCCTCATTTAAT-----666
Db      :::::::::::
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
Db      :::::::::::
QY 667 -----GACITC 672
Db      :::::::::::
QY 201 ProArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIle 220
Db      :::::::::::
QY 673 ACTCGGGTTGTTGGTGGAGAGATGCCAACCCAGGTCAATTCCTTGGCAGGTTGTTTG 732
Db      :::::::::::
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThr 240
Db      :::::::::::
QY 733 AATGTGTAAGTTGATGTCATTTCTGTGAGGCTCTATCGTTAATGAAATAATGATGTAAC 792
Db      :::::::::::
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
Db      :::::::::::
QY 793 GCTGCCCACTGTGTGAAACTGGTGTAAAT-----ACAGTTGTCCGAGGTGAA 843
Db      :::::::::::
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIle-----276
Db      :::::::::::
QY 844 CATATATTGAGGAGACAGAACATACAGAGCAAAAGCGAATGTGATTCGATTTATTCCT 903
Db      :::::::::::
QY 277 ---IleGluPheAsnProMetTyrProLysAspAsn---AspIleAlaLeuMetLysLeu 294
Db      :::::::::::
QY 904 CACACAACTACAATGCGACTATTAAATAGTACAAACCATGACATGACATTCCTTCTGGA 963
Db      :::::::::::
QY 295 GlnPheProLeuThrPheSerGlyThrValArgProIleCysLeuProPheAspGlu 314
Db      :::::::::::
QY 964 GACGAACCCCTAGTGCTAAACAGCTACCTTACACCTATTGCAATGCT-----GACA 1017
Db      :::::::::::
QY 315 GluLeuThr-----ProAlaThrProLeuTrpIleIleGlyTrpGly-----328
Db      :::::::::::
QY 1018 GAATACACAGACATCTTCTCAATTTGGATCTGGCTATGTAAGTGGCTGGGGAAGATC 1077
Db      :::::::::::
QY 329 PheThrLysGlnAsnGlyGlyLysMetSerAspIleLeuLeuGlnAlaSerValGlnVal 348
Db      :::::::::::
QY 1078 TTCCACAAA-----GGGAGATCAGCTTTAGTTCTTCTAGTACCTTGAAGTTCCACT 1128
Db      :::::::::::
QY 349 IleAspSerThrArgCysAsnAlaAspAspAlaTyrGlnGlyGluValThrGluLysMet 368
Db      :::::::::::
QY 1129 GTTGACCGAGCCACATGCTCTCGATCTACAAAAGTTC-----ACCATCTATAACAACATG 1182
Db      :::::::::::
QY 369 MetCysAlaGlyIleProGluGlyGlyValAspThrCysGlnGlyAspSerGlyGlyPro 388
Db      :::::::::::
```

R:Keime, S.; Adham, I.M.; Engel, W.
 Eur. J. Biochem. 190, 195-200, 1990
 A:Title: Nucleotide sequence and exon-intron organization of the human proacrosin gene.
 A:Reference number: S11674; MUID:90306003; PMID:2114285
 A:Accession: S11674
 A:Molecule type: DNA
 A:Residues: 1-421 <KE1>
 A:Cross-references: UNIPROT:P10323; UNIPARC:UPI000012539D; EMBL:X54017; NID:g35582; PIDN
 A:Note: the authors translated the codon AGG for residue 64 as Thr and CTG for residue 2
 R:Valquez-Levin, M.H.; Revetos, J.; Gordon, J.W.
 Eur. J. Biochem. 207, 23-26, 1992
 A:Title: Molecular cloning, sequencing and restriction mapping of the genomic sequence of
 A:Reference number: S23499; MUID:92331659; PMID:1628652
 A:Accession: S23499
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-421 <VAZ>
 A:Cross-references: UNIPARC:UPI000012539D; EMBL:M77378
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1992
 R:Keime, S.
 submitted to the EMBL Data Library, December 1989
 A:Reference number: S12063
 A:Accession: S12063
 A:Molecule type: DNA
 A:Residues: 1-225, 'R', 227-421 <KE12>
 A:Cross-references: UNIPARC:UPI0000172B0D; EMBL:X54017
 R:Adham, I.M.; Klemm, U.; Maier, W.M.; Engel, W.
 Hum. Genet. 84, 125-128, 1990
 A:Title: Molecular cloning of human preproacrosin cDNA.
 A:Reference number: A61022; MUID:90128988; PMID:2298447
 A:Accession: A61022
 A:Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-63, 'T', 65-225, 'V', 227-267, 'R', 269-421 <ADH>
 A:Cross-references: UNIPARC:UPI0000172B0E
 R:Baba, T.; Watanabe, K.; Kashiwabara, S.I.; Arai, Y.
 FEBS Lett. 244, 296-300, 1989
 A:Title: Primary structure of human proacrosin deduced from its cDNA sequence.
 A:Reference number: S03330; MUID:89153568; PMID:2493394
 A:Accession: S03330
 A:Molecule type: mRNA
 A:Residues: 1-63, 'T', 65-119, 'V', 121-165, 'L', 167-267, 'R', 269-344, 'R', 346-421 <BAB>
 A:Cross-references: UNIPARC:UPI0000167BCA; EMBL:X00970; NID:g28325; PIDN:CAA68784.1; PID
 C:Genetics:
 A:Gene: GDB:ACR
 A:Cross-references: GDB:119645; OMIM:102480
 A:Map position: 22q13-22qter
 A:Introns: 26/2; 94/2; 189/2; 237/3
 C:Superfamily: acrosin; trypsin homology
 C:Keywords: glycoprotein; hydrolase; serine proteinase; sperm
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-421/Product: acrosin #status predicted <MAT>
 F:20-421/Product: acrosin light chain #status predicted <LCH>
 F:43-421/Product: acrosin heavy chain #status predicted <HCH>
 F:43-285/Domain: trypsin homology <TRY>
 F:302-379/Region: proline-rich
 F:22,210/Binding site: carbonylrate (Asn) (covalent) #status predicted
 F:25-154/Disulfide bonds: #status predicted
 F:29-162/Disulfide bonds: #status predicted
 F:73-89/Disulfide bonds: #status predicted
 F:88,142,240/Active site: His, Asp, Ser #status predicted
 F:177-246/Disulfide bonds: #status predicted
 F:209-225/Disulfide bonds: #status predicted
 F:236-266/Disulfide bonds: #status predicted

Query Match 20.5%; Score 481; DB 1; Length 421;
 Best Local Similarity 41.0%; Pred. No. 4,5e-32;
 Matches 100; Conservative 38; Mismatches 88; Indels 18; Gaps 5;

QY 202 RYVGGEEASVDSMPWQVSIQY-----DKQHVCGGSLDPHWVLTAAHCFRKHTDVPNNKV 256
 DB 42 RIVGGKAAQHGAQPMWVSIQIFRYSNRHYTCGSLNSRWVLTAAHCFVGRNNVHWRL 101

QY 257 RAGSDKL---GSFPSLA-----VAKIIIEFNPMPKNDIALMKIQFPLETSGTVRPI 307
 DB 102 VFGAKETTYGNNKRVKAPLQERYEKIIIEKYNASATEGDIALVEITPTISGCRFIGPG 161
 QY 308 CLPFFDELTLPAT-PLMIIIGMFTKONGGKMSDILLQASQVVIDSTRCNADDAQGEVTE 366
 DB 162 CLPFFKAGLPRGSQSCWVAGWGYEEKAPRPSSTILMEARVDLIDLCSNTQWNGRVQ 221
 QY 367 KMGAGIPEGVDPTCGQDSGGPLMYQSDQ---MHVIGVSWGYCGGPGPTPGVYTKVSAY 423
 DB 222 TNVCAGIFVQKIDTCQDSGGPLMKCKDSKESAVYVVGITSMGVGCALAKRPGITATMPY 281
 QY 424 LNMW 427
 DB 282 LNMW 285

Search completed: December 5, 2005, 13:53:51
 Job time : 41 secs

Db 169 PRPSVLLAEARVELLDLCNSTQWYNGVMSSTNWCAGYEGKIDPTCGSGGFLMCRDN 248
 Qy 393 -SDQWNVGVISWGCGGPGSTPGYTKVSAIYLM 427
 Db 249 ANSPFVVVGIITSMGVCARAKRPGIYTAWTYLDWI 284

RESULT 13

KOHUP

plasma kallikrein (EC 3.4.21.34) precursor - human

N/Alternate names: kininogenin; plasma prekallikrein

C/Species: Homo sapiens (man)

C/Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 05-Oct-2004

C/Accession: A00921; A37939

R/Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.

Biochemistry 25, 2410-2417, 1986

A/Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four

A/Reference number: A00921; PMID:86243359; PMID:3521732

A/Accession: A00921

A/Molecule type: mRNA

A/Residues: 1-638 <CHU>

A/Cross-references: UNIPROT:P03952; UNIPARC:UPI00000088AE; GB:M13143; NID:G190262; PIDN:

Biochemistry 30, 2050-2056, 1991

A/Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of

A/Reference number: A37939; PMID:91152016; PMID:1998666

A/Accession: A37939

A/Molecule type: protein

A/Residues: 20-27,40-46, 'X', 48, 'H', 50, 'X', 52-70, 'H', 75-76, 'X', 78-80, 103-113, 131-140, 141-

525,538-551,562, 'X', 564-567,573, 'X', 575-576,578-583, 'X', 585,592-604 <MCM>

A/Cross-references: UNIPARC:UPI00000172B4; UNIPARC:UPI00000172B4B; UNIPARC:UPI00000172B4C; UNIPARC:UPI00000172B4D;

B49; UNIPARC:UPI00000172B4A; UNIPARC:UPI00000172B5; UNIPARC:UPI00000172B5B; UNIPARC:UPI00000172B5C; UNIPARC:UPI00000172B5D;

C/Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w

are linked by one or more disulfide bonds.

C/Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li

C/Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r

C/Genetics:

A/Gene: GDB:KLX3

A/Cross-references: GDB:127575; OMIM:229000

A/Map position: 4q35-4q35

C/Superfamily: plasma kallikrein/coagulation factor XI; trypsin homology

C/Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla

F/1-19/Domain: signal sequence #status predicted <SIG>

F/20-638/Product: plasma kallikrein #status predicted <MAT>

F/20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>

F/110-199/Domain: apple repeat <AP1>

F/200-289/Domain: apple repeat <AP2>

F/291-380/Domain: apple repeat <AP3>

F/391-638/Domain: plasma kallikrein light chain #status predicted <LCH>

F/391-638/Domain: trypsin homology <TRY>

F/21-104,47-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383

F/127-308,396,453,494/binding site: carboxylate (Asn) (covalent) #status experimental

F/318-347,349-345/disulfide bonds: #status predicted

F/390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted

F/434,483,578/Active site: Arg, Asp, Ser #status predicted

Query Match 21.0%; Score 491.5; DB 1; Length 638;

Best Local Similarity 42.2%; Pred. No. 1e-32; Indels 16; Gaps 8;

Matches 100; Conservative 47; Mismatches 74; Indels 16; Gaps 8;

Qy 202 RVVGEASVDSMPQVSIQYD--KQHVCGSILDPHVLTAHCF--RQHTVFNWV 256
 Db 390 RIVGTSNMGWMPQVSLQVLTQGRHLGSGSLIGHQWVLTAAHCFDGLQDV--WRI 447
 Qy 257 RAG----SDKGSFSLVAKIIIEFNPMYRKNDIALMLQLPPLTFSGVTRDCLPFF 312
 Db 448 YSIIINLSIDITKDPFQIKETIIHQNYKVSQGNHDIALLQLQPLNVTBQKRICLPK 507

Qy 313 DEELTPATPLMIIIGCFYTKONGCKRMSDILQASVOVIDSTRCNADAYQG-EVTEKMGCA 371
 Db 508 GDRSTIYTKCWYMGWGFSSHEK-GEIONILQKXNIPLVTEEC--QKRQDYKIIITRMCA 564
 Qy 372 GPEGVDTCCGDSGGPLMYQ--SDQWNVGVISWGCGGPGSTPGYTKVSAIYLM 427
 Db 565 GYEGSGKDCCKDSGGPLVCYKINMKRLVGIITSMGVCARRBQPGVYTKVAEYMDWI 621

RESULT 14

JEO315

low-density lipoprotein receptor-related protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C/Accession: JEO315

R/Tomita, Y.; Kim, D.; Magocsi, K.; Fujino, T.; Yamamoto, T.T.

J. Biochem. 124, 784-789, 1998

A/Title: A novel low-density lipoprotein receptor-related protein with type II membrane

A/Reference number: JEO315; PMID:98429596; PMID:9756624

A/Accession: JEO315

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-1113 <TOM>

A/Cross-references: UNIPROT:Q92319; UNIPARC:UPI0000029874; DDBJ:AB013874; NID:G3869144;

C/Superfamily: mouse low-density lipoprotein receptor-related protein; LDL receptor li

F/337-372/Domain: LDL receptor ligand-binding repeat homology <LDL1>

F/410-445/Domain: LDL receptor ligand-binding repeat homology <LDL2>

F/447-482/Domain: LDL receptor ligand-binding repeat homology <LDL3>

F/648-682/Domain: LDL receptor ligand-binding repeat homology <LDL4>

F/723-757/Domain: LDL receptor ligand-binding repeat homology <LDL5>

F/869-1097/Domain: LDL receptor ligand-binding repeat homology <LDL7>

Query Match 20.9%; Score 488.5; DB 2; Length 1113;

Best Local Similarity 33.2%; Pred. No. 3.5e-32;

Matches 130; Conservative 52; Mismatches 148; Indels 61; Gaps 17;

Qy 69 IPRQLDGLDCLPLGDEEHCVKSPFEGPAVAVRLSKRSTQVLDATGNMFSACFDN 128
 Db 736 VPRDMLCDGWDCSDSDSEMGCVTLKNG-----NSSLLTYHKAHEH--HYCAGD 785
 Qy 129 FTEALFACRKMGYSKPTFRAVEIGPODDLVETENSQELRM---RNSG----- 179
 Db 786 WRETLSQLACKQWGLGEPSTKLI---PQV-----EQQWLRILYPMWENLNGSTLQ 834
 Qy 180 -----PCLSGSLVSLHCLA--CGK---SLKTPRVVGEPAVSDVSPWQVSTQYDKQ-HV 227
 Db 835 LLYVRHSCPSRSEISLCSKQDGRPPAAMNRKIIIGKTSRGRPMWQCSIQSEPSGHI 894
 Qy 228 CGGSILDPHVLTAHCFRKHNDVFNWVYRAGSDKL--GSFSLAVAKIIIEFNPMY 284
 Db 895 CGCVLLAKKVVLTVAHCFEGREBADVAKVFGINNLDHPSGFGQTRFVKTILL--HPRS 952
 Qy 285 K---DNDIADMLKQLPPLTFSGVTRPICFPPEBELTPAPVLTIGGFPKQNGKMSDIL 341
 Db 953 RAYVVDIVSVELSDINETSIVRPVCLPSPEYLEPDYCYITGW--HMGNTKMPFL 1009
 Qy 342 LQASVOVIDSTRCNADAY--QGEVTEKMGCAIGEGVDTCCGDSGGPLMYQ--SDQW 397
 Db 1010 QEGEVRIILPLEQ---QSYFDMKTTINRMICAGYESGTYDSCMGDSGGLVCEERPGQWT 1066
 Qy 398 VVGVISWGCGGPGSTPGYTKVSAIYLM 427
 Db 1067 LRELISWGSVCFPSKVLGPGVSNVSPVFGWI 1097

RESULT 15

S11674

acrosin (EC 3.4.21.10) precursor - human

C/Species: Homo sapiens (man)

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: S11674; S23499; S12063; S03330

A:Residues: 1-68 <KLE>
A:Cross-references: UNIPARC:UPI000016CBCE; GB:X52466; NID:g49857; PIDN:CAA36704.1; PID:g
C:Superfamily: acrosin; trypsin homology
C:Keywords: glycoprotein; hydrolase; serine proteinase
F:39-283/Domains: trypsin homology <TRY>
F:18_208/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:21-151/Disulfide bonds: #status predicted
F:25-158/Disulfide bonds: #status predicted
F:70-86/Disulfide bonds: #status predicted
F:85_139_238/Active site: His, Asp, Ser #status predicted
F:175-244/Disulfide bonds: #status predicted
F:207-223/Disulfide bonds: #status predicted
F:234-264/Disulfide bonds: #status predicted

Query Match 21.5% Score 503; DB 2; Length 418;
Best Local Similarity 39.1%; Pred.No. 6.7e-34;
Matches 108; Conservative 36; Mismatches 90; Indels 42; Gaps 7;

OY 179 GPLSGSLVSLHCLACGSLKT-----PRVGGEEASVSWMQVSIQY-----DKQ 225
||| :||| :||| :||| :||| :
DB 23 GPC-----GLRTRONSOAGTRIVSQSAHVAMPMWSLQIFSHNSRRY 67

OY 226 HVCGSILDPHWVLTAHCFRKHTDVFNMK-----VRASDLGSP--SLAVAKII 276
||||| :||| :||| :||| :||| :
DB 68 HACGSGLNSHMVLTAHCDFNKKKVYDMRLVFGAQETIEYRKNPKVEPQEERYVQKI 127

OY 277 IEFNPMPKNDIALMKQLFPFLTFSGITVRPICLPFDEELTPATP--LMIIGWFTKONG 334
||| :||| :||| :||| :||| :
DB 128 HEKNVVVTEGNDIALTKTPTVTGCNFIQPCCLPHFKAPGRPKIPIHCYVIGWGIARE 187

OY 335 GRMSDILLQASVOVIDSTRCNADDAYOGEVTERKMKGACIPRGGVDTCCGDSGGPLATQSD 394
: ||| :||| :||| :||| :||| :
DB 188 PRPSVLTAEARVDLDIDLCLNSTQWYNNGRVSTNVCAYPEGKIDTCGDSGGPLMCRDN 247

OY 395 ---QMHNVIYSWGVCGCCGPSTPGYTTSKVAYLMI 427
: ||| :||| :||| :||| :||| :
DB 248 ARQPFVVVGITSWGVCARAKRPGYVYATWDYLDWI 283

RESULT 11
S18407
acrosin (EC 3.4.21.10) precursor - rat
N:Contents: proactosin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S18407; S30037; A56620
R:Kleemm, U.; Flake, A.; Engel, W.
Biochim. Biophys. Acta 1090, 270-272, 1991
A>Title: Rat sperm acrosin: cDNA sequence, derived primary structure and phylogenetic o
A:Reference number: S18407; MUID:92031708; PMID:1932123
A:Accession: S18407
A:Molecule type: mRNA
A:Residues: 1-437 <KLE>
A:Cross-references: UNIPROT:P29293; UNIPARC:UPI00001253A0; EMBL:X59254
R:Kleemm, U.; Flake, A.; Engel, W.
Submitted to the EMBL Data Library, April 1991
A:Reference number: S30037
A:Accession: S30037
A:Molecule type: mRNA
A:Residues: 1-254, 'LCDR', 259, 'DHBL', 264, 'GRIC', 269-437 <KLE2>
A:Cross-references: UNIPARC:UPI0000170B23; EMBL:X59254; NID:g57282; PIDD:CAA41947.1; PIR
R:Kremling, H.; Flake, A.; Adam, I.M.; Radtke, J.; Engel, W.
DNA Seq. 2, 57-60, 1991
A>Title: Exon-intron structure and nucleotide sequence of the rat proactosin gene.
A:Reference number: A56620; MUID:92199245; PMID:1802037
A:Accession: A56620
A>Status: preliminary
A:Molecule type: DNA; mRNA
A:Residues: 1-254, 'LCDR', 259, 'DHBL', 264, 'GRIC', 269-437 <KRE>
A:Cross-references: UNIPARC:UPI0000170B23
A>Note: sequence modified after extraction from NCBI backbone
A:Note: sequence extracted from NCBI backbone (NCBIN:89436, NCBIN:89439, NCBIN:89447, NC
C:Superfamily: acrosin; trypsin homology

C/Keywords: glycoprotein; hydrolase; serine proteinase; sperin; zymogen
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-43/Domain: acrosin light chain #status predicted <LCH>
F:43-286/Domain: trypsin homology <TRY>
F:44-437/Domain: acrosin heavy chain #status predicted <HCH>
F:22-211/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:25-155-29-163/Disulfide bonds: #status predicted
F:74-90/Disulfide bonds: #status predicted
F:89-143-241/Active site: His, Asp, Ser #status predicted
F:178-247/Disulfide bonds: #status predicted
F:210-226/Disulfide bonds: #status predicted
F:237-267/Disulfide bonds: #status predicted

Query Match 21.4%; Score 501.5; DB 2; Length 437;
Best Local Similarity 42.7%; Pred. No. 9,4e-34;
Matches 105; Conservative 35; Mismatches 85; Indels 21; Gaps 6;

QY 202 RVVGGEEASVDSMPWQVSIQY-----DKQVCGSTLDPHWVLTAAHCFKHTDVFMMK 255
||:||||:|
DB 42 RIVGGOTSRMAMPWVNSLQIFTSNRSRYAACGSLNSHWLTAAHCFDKKKVYDNR 101
||:||||:|
QY 256 -----VAGSDKLGSFSLA--VAKIIIEFNMYKPNKDIALMKLOFPLTFSGTVP 306
||:||||:|
DB 102 LVFGAHEIYGNKKPVEKFOQERYVQKIVHEKKNATBEGNDIALKTPVPTCGDFGP 161
||:||||:|
QY 307 ICLPFDEELTPATP--LWIIIGFTKONGKMSDILLQASVQVIDSTRCNADAYQGEV 364
||:||||:|
DB 142 GCLPHF-KSGPRLPHTCYVTGWGVIKDNAPRSPVIMEARVDLIDLCLNSTQWYNGRV 220
||:||||:|
QY 365 TEKMCAGIPEGGVPTCCGDSGGPLMYQSDQ--MHNVGYVSKGVCGGSPSTPGYTKVS 421
||:||||:|
DB 221 TSTNVCAGPEGKITDCCGDSGGPLMCRDTRRQPFVIGITSWGVCARAKRPGVYATW 280
||:||||:|
QY 422 AYLNWI 427
||:||||
DB 281 DYLDWI 286
||:||||

RESULT 12
S29599
acrosin (EC 3.4.21.10) precursor - guinea pig (fragment)
C/Species: Cavia porcellus (guinea pig)
C/Date: 22-Nov-1993 #sequence_revansion 10-Nov-1995 #text_change 09-Jul-2004
C/Accession: S29599
R/Gerton, G.L.; Hoff, H.B.; Baba, T.
submitted to the EMBL Data Library, May 1992
A/Description: The amino acid sequence of guinea pig proacrosin deduced from its cDNA se-
A/Reference number: S29599
A/Accession: S29599
A/Molecule type: mRNA
A/Residues: 1-421 <GER>
A/Cross-references: UNIPROT:O60491; UNIPARC:UPI00000E5C67; EMBL:Z12153; NID:G49559; PIDN
C/Superfamily: acrosin; trypsin homology
C/Keywords: glycoprotein; hydrolase; serine proteinase
F:41-284/Domain: trypsin homology <TRY>

Query Match 21.1%; Score 493.5; DB 2; Length 421;
Best Local Similarity 38.0%; Pred. No. 4.1e-33;
Matches 105; Conservative 44; Mismatches 84; Indels 43; Gaps 8;

QY 179 GPC-----LGSGLVSLHCLACGKSLKTRPVVGGEEASVDSMPWQVSIQY-----DKQ 225
||:||||:|
DB 25 GPCGAFRQNLQGSV-----RIIGGQFAQPGAMPWVNSLQIFMAHNNRY 69
||:||||:|
QY 226 HVCGSILDPHWVLTAAHCFKHTDVFNMKVRAAGDKL--GSPSLA-----VAKIII 276
||:||||:|
DB 70 HACGGILNSHWLTAAHCFDSKKKYDMLVFGEAEIEYGNKKVRAAPLOERYVEKTVI 129
||:||||:|
QY 277 IEFNPMYKPNKDIALMKLOFPLTFSGTVPRLCLPFDEELTPATP--LWIIIGFTKONG 334
||:||||:|
DB 130 HEKKNATBEGNDIALKTPVPGCGPLTPTF-RAGSPKIPQTCYVAGWGIIRREA 188
||:||||:|
QY 335 GKMSDILLQASVQVIDSTRCNADAYQGEVTEKMCAGIPEGGVPTCCGDSGGPLMYQ-- 392

F:21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,322
F:127,215,308,396,434/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:434,483,578/Active site: His, Asp, Ser #status predicted

Query Match

21.7%; Score 508; DB 1; Length 638;

Best Local Similarity 41.0%; Pred. No. 4,3e-34;

Matches 112; Conservative 45; Mismatches 82; Indels 34; Gaps 10;

QY 174 MRNSGPCLSGSLVSHCLACGKSLKTP-----RVNGGEASVDSMPQVSIQ---Y 222

DB 364 MQSSSG-----YSRLCKLVDSPTCTKINARIVGCTNASSIGEMWQVSLQVLY 413

QY 223 DKQHNCGSLTDHMTLTAHCFR--KHTDVFMKTRAG---SDKLGSFPLAIAVKI 276

DB 414 SQTHLCGGSLIGRWVLTAAHCPDGIYPDV--WRIYGGTILSLSEITKETSRIKEL 471

QY 277 IEFNPMVKPNDIATLMLKLPPLTFSGTVPRICLPFDEELTPATPLMIIGMGTQNGK 336

DB 472 HQRYKSEGVYDALIKLQPLNLTPEQKPCILPSKADNTITTCNVGTGMYTKEQ-GB 530

QY 337 MSDILLQASVQVDSITRCNADDAVQGEVTEK--MNCAGIPEGVDTCCGDSGGLMTQ-SD 394

DB 531 TQNILOKATRIPLVNEEC--QKRYRDYVINKQMICAGYKCGTACKGDSGGLVCKHG 588

QY 395 QMHWGVGVGVCYCGSPSTPGVYTKYSAYIANMI 427

DB 589 RMQLVGITSMWEGCGRRKDPGVYTKVSEYWDWI 621

RESULT 8

JX0172

acrosin (EC 3.4.21.10) precursor form 3 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C/Accession: JX0172; JX0138

R:Watanabe, K.; Baba, T.; Kashiwabara, S.; Okamoto, A.; Arai, Y.

J. Biochem. 109, 828-833, 1991

A>Title: Structure and organization of the mouse acrosin gene.

A/Reference number: JX0172; MUID:92041732; PMID:1939002

A/Accession: JX0172

A/Molecule type: DNA

A/Residues: 1-436 <MAT>

A/Cross-references: UNIPROT:P23578; UNIPARC:UPI000021F79; GB:S66245; NID:9238706; PIDN:

R:Kashiwabara, S.; Baba, T.; Takada, M.; Watanabe, K.; Yano, Y.; Arai, Y.

J. Biochem. 108, 785-791, 1990

A>Title: Primary structure of mouse preacrosin deduced from the cDNA sequence and its ge

A/Reference number: JX0138; MUID:91185335; PMID:2127931

A/Accession: JX0138

A/Molecule type: mRNA

A/Residues: 4-436 <KAS>

A/Cross-references: UNIPARC:UPI000016C8D; GB:D00754; NID:9220322; PIDN:BAA00651.1; PID:

C/Comment: Acrosin is an acrosomal protease that plays an important role in the initial

e-binding activity.

C/Genetics: 26/2; 95/2; 190/1; 238/3

A/Introns: 26/2; 95/2; 190/1; 238/3

C/Superfamily: acrosin; trypsin homology

C/Keywords: glycoprotein; hydrolase; serine protease; sperm

F:1-19/Domains: signal sequence #status predicted <SIG>

F:20-42/Product: acrosin light chain #status predicted <ALC>

F:43-321/Product: acrosin heavy chain #status predicted <AHC>

F:43-286/Domains: trypsin homology <TRY>

F:22.211/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:25-155,29-162,74-90,178-226,237-267/Divalent bonds: #status predicted

F:89,143,241/Active site: His, Asp, Ser #status predicted

F:235/Binding site: substrate (Asp) #status predicted

Query Match 21.6%; Score 505; DB 2; Length 436;

Best Local Similarity 40.9%; Pred. No. 4.8e-34;

Matches 106; Conservative 37; Mismatches 90; Indels 26; Gaps 7;

QY 194 CG-----KSLKTPRVYVGEASVDSMPQVSIQ-----DKQHNCGSLTDHMTLTA 242

DB 29 CGLRFRNSQAGTRIVSGSQSLTGAMPWVSLQIFTSNRSRYHACGSLNSHWLTA 88

QY 243 HCFRKHITDVFNK-----VRAGSDKLGSFPLA--YAKIIIEFNPMVKPNDIATL 293

DB 89 HCFDNKKKYDWRVLTGAQAEIYGRNKPVEKQGRYVQKIVIHKKYNVTEGNDIALK 148

QY 294 LQPLTFSGTVPRICLPFDEELTPATP--LMIIGMGTQNGKMSDILLQASVQVDS 351

DB 149 ITPPVTCGNFIPGCPCLPHF--KAGPQIPIHTCVVTGVIKKAAPRPSPLMEARVLD 207

QY 352 TRCNADDAVQGEVTEKMKACGIPGGVDTCCGDSGGLMTQSD---QMHWGVGVGVC 408

DB 208 DLCSNIGWNGRVISTNVCAGIPBKGIDTCQDSGGLMCRDNDSPFVYVGTSMWGC 267

QY 409 GGPSTPGVYTKVSAIYANMI 427

DB 268 ARAKRPGVYATATWYLDWI 286

RESULT 9

T30337

polyprotein - African clawed frog

C/Species: Xenopus laevis (African clawed frog)

C/Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C/Accession: T30337

R:Yang, J.C.; Lindsey, L.L.; Hedrick, J.L.

submitted to the EMBL Data Library, March 1998

A/Description: cDNA cloning of ovocytinase, a chymotrypsin-like protease released from X

A/Reference number: Z20829

A/Accession: T30337

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Residues: 1-1524 <YAN>

A/Cross-references: UNIPROT:Q91674; UNIPARC:UPI0000FBA76; EMBL:U81290; NID:92581640; P

C/Superfamily: trypsin related polyprotein; trypsin homology

Query Match

21.5%; Score 504.5; DB 2; Length 1524;

Best Local Similarity 44.0%; Pred. No. 2.4e-33;

Matches 109; Conservative 38; Mismatches 86; Indels 15; Gaps 8;

QY 194 CGKSLKT-----PRVNGGEASVDSMPQVSIQDKQVCGSLTDHMTLTAHCFPKH 248

DB 570 CGMAWTRKMWLPRIYVGEASVDSMPQVSIQFPLRTHCEGALISPMWITLTAHCFIAA 629

QY 249 TDVFNWKTFRAGS--DKLGSFPLAIAVAKIIIEFNPMYK---DNIDIALMLKLPPLTFSGTV 304

DB 630 EBSY-WYIYAGDHRMNLNESTEQIRNITIRIHNYNSETYDNDIALYLSEPLDINPV 688

QY 305 RPICLPFDEELTPATPLMIIGMGTQNGKMSDILLQASVQVDSITRCNADDAVQGEV 364

DB 689 RPYCLPEDEEVLTPASVCVWGTADGQPALG-LQQLPLIDSLICNT--SYSGEL 746

QY 365 TEKMKACGIPBG--GVDTCCGDSGGLMTQSD---DMHWGVGVGVCSPSTPGVYTKVS 421

DB 747 TDHMLCAFPSPSKEXKDCQDSGGLVYQNKKEQFSIYGLVSWGCGCRVSPGVYTKVR 806

QY 422 AYLMWYIN 429

DB 807 LFTWYIQN 814

RESULT 10

A37344

acrosin (EC 3.4.21.10) precursor form 1 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 22-Jun-1999

C/Accession: A37344

R:Klemm, U.; Meier, W.M.; Tsasoudou, S.; Adam, I.M.; Willison, K.; Engel, W.

A/Title: Mouse preacrosin: cDNA sequence, primary structure and postmeiotic expressi

A/Reference number: A37344; MUID:90255839; PMID:2111255

A/Accession: A37344

A/Status: preliminary

A/Molecule type: mRNA

F:801-1030/Domain: trypsin homology <TRY>
 F:116,147,170,194,233,263,264,404,456,486,519,550,646,698,722,741,762,864,903,965/Bindin
 F:788-912,826-942,957-972,983-1011/Dissulfide bonds: #status predicted
 F:841,892,987/Active site: His, Asp, Ser #status predicted

Query Match 24.4%; Score 571; DB 1; Length 1035;
 Best Local Similarity 35.0%; Pred. No. 4,6e-39;

Matches 133; Conservative 67; Mismatches 138; Indels 42; Gaps 15;

QY 69 IPRKOLCDDELDCPLGEDEEHCVKSPPEGPAAVAVRLSKRSTLOVDSATGMFSAFCPN 128
 DB 672 IPLVNLCDGFPHCKGSDAHCVRLF-NTTDSGLVQRR--IQSI-----MHYACAE 722
 QY 129 FTEALAEATACRQMGY----SSKPTFAVEIGPDODLVE-----ITENSQELMRNSSG 180
 DB 723 WTTQISDVQCQLGLGTGNSVPTF-STGGPYVNLNTAPNSGLITPEQQ----- 772
 QY 181 CLSGSLVSLHLC--LACGSKLKT---PRVYGGEEASVDSMPQVSIQYDKHVCSSIL 234
 DB 773 CLEDSLILLCQCNVKSCKKLVTQEVSPKIVGSDSREGAMPVVALYPDDQQVCGASIVS 832
 QY 235 PHWVLTAAHC-PRKHIDVENMKVRAGSDKLSF--PSLAVAKIIIEFPMV---PKMD 288
 DB 833 RMVLSAAHCVCYGRNMBESKMAVGLHMASNLTSFOITRLIDQIVINPHNKRKXND 892
 QY 289 IALMKLQPLPFGSVTRPICLPFDELTTPATPLMIIGFTKONGKMSDILLQASVQV 348
 DB 893 IAMHLEMKVNTDYIQPLCLPEBNQVFPGRICSLAGGALIIYQ-GSTADVLQEDVPL 951
 QY 349 IDSTCNADDAVQGEVTEKMCAGIPGGVDTTCQSDSGPLMYQ--SDQMHVGVISWG 407
 DB 952 ISNECKQ-QQMEPVYITENMVCAGYEAQGVDSQQGDSGGLMCOENRWLLAGVTSFG 1010
 QY 408 CGGSPGPGYTKVSAVLMWI 427
 DB 1011 CALPRPGYARVPRTEWI 1030

RESULT 4

S33777
 hepsin (EC 3.4.21.-) - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 06-Jan-1995 #sequence _revision 06-Jan-1995 #text_change 05-Oct-2004
 C:Accession: S33777; S32013
 R:Farley, D.; Raymond, F.; Nick, H.
 Biochim. Biophys. Acta 1173, 350-352, 1993
 A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteinase.
 A:Reference number: S33777; MUID:93305733; PMID:8318546
 A:Accession: S33777
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-416 <FAR>
 A:Cross-references: UNIPROT:Q05511; UNIPARC:UP100004ECM9; EMBL:X70900; NID:957928; PIDN
 C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein
 F:12-44/Domain: transmembrane #status predicted <TM>
 F:162-399/Domain: trypsin homology <TRY>
 F:187-203,280-358,321-337,348-380/Dissulfide bonds: #status predicted
 F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 24.3%; Score 568.5; DB 1; Length 416;
 Best Local Similarity 30.5%; Pred. No. 2.5e-39;
 Matches 137; Conservative 64; Mismatches 155; Indels 93; Gaps 11;

QY 19 KPRIMEPRFKGIIITALLSLAIIIVVLIKVIYLDKTYFLCGOPHFIKPKOLCDGE 78
 DB 14 RPKVAALT---VGLILFLIGASMAIYITILLR----- 44
 QY 79 LDCPLGEDEEHCVKSPPEGPAAVAVRLSKRSTLOVDSATGMFSAFCPNFTEALATAC 138
 DB 45 -----SQQE-----PLYQVQLSPGDSRLVLVDKTEGTWRLLCSSRSNARAVAGIGC 89
 QY 139 RQMGY-----SSKPTFAVEIG---PDODLVEITENSQELMRNSS 178

DB 90 EEMGLRALANSELDVRTAGANGTSGFECVGGGLPLAQRLLDVISVCD----- 138
 QY 179 GPCLSGSLVSLHCLACG-KSLKTPVVGEEASVDSMPQVSIQYDKHVCSSILDPHW 237
 DB 139 --CPRGRLTATQCQCGCRKLPLVDRIVGGQSSLSKRWMPVSLRDKDTHLCCGSLSGDW 196
 QY 238 VLTAAHCERKATDVN-WKVRAGSDKLSFPSLAVALKIIIEFNMYP-----KOND 288
 DB 197 VLTAAHCPEERNRVLRSRNVFAGAVARTSPHVGQVAVIYHGSLVLPFRDPTIDENSND 256
 QY 289 IALMKLQPLPFGSVTRPICLPFDELTTPATPLMIIGFTKONGKMSDILLQASVQV 348
 DB 257 IALVHLSSSLPLETEYIQVCLPAAGQALVDKVCVTVMGNT-QPYGQAVVLDQEARVPI 315
 QY 349 IDSTCNADDAVQGEVTEKMCAGIPGGVDTTCQSDSGPLMYQ-----SDQMHVGVIS 403
 DB 316 ISNEVCNSPDYGNQIKRKMFCAGIPGGIDTCQSDSGHFFCEDRISGTSRMRCLGIVS 375
 QY 404 WGYGCGSPGPGYTKVSAVLMWIYVWK 432
 DB 376 WGTGALARKPGVYTKVIDPREWIFQAIK 404

RESULT 5

A53663
 enteropeptidase (EC 3.4.21.9) precursor [validated] - pig
 N:Alternate names: enterokinase
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 07-Oct-1994 #sequence _revision 09-Aug-1996 #text_change 28-Apr-2003
 C:Accession: A53663
 R:Matsumura, M.; Ichinose, M.; Yahagi, N.; Kakel, N.; Tsukada, S.; Miki, K.; Kurokawa,
 J. Biol. Chem. 269, 19976-19982, 1994
 A:Title: Structural characterization of porcine enteropeptidase.
 A:Reference number: A53663; MUID:94327548; PMID:8051081
 A:Accession: A53663
 A:Molecule type: mRNA
 A:Residues: 1-1034 <MAT>
 A:Cross-references: UNIPARC:UP10000172B0C; GB:D30799; NID:9505122; PIDN:BA06459.1; PID
 A:Note: parts of this sequence, including the amino ends of three chains isolated from t
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
 C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
 C:ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involv
 C:Function:

A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolyase cascade
 C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding rep
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
 F:22-38/Domain: transmembrane #status predicted <TM>
 F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>
 F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>
 F:119-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:357-519/Domain: MAM homology <MAM>
 F:541-646/Domain: C1r/C1s repeat homology <C1R>
 F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC
 F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>
 F:800-1029/Domain: trypsin homology <TRY>
 F:116,147,170,194,283,343,350,403,455,485,519,549,645,697,701,721,740,761,804,863,902,96
 F:787-911,825-941,925-992,956-971,982-1010/Dissulfide bonds: #status predicted
 F:840,891,986/Active site: His, Asp, Ser #status predicted

Query Match 23.6%; Score 552; DB 1; Length 1034;
 Best Local Similarity 34.2%; Pred. No. 1.7e-37;
 Matches 127; Conservative 69; Mismatches 141; Indels 34; Gaps 14;

QY 74 LDCPLGEDEEHCVKSPPEGPAAVAVRLSKRSTLOVDSATGMFSAFCPNFTEAL 133
 DB 676 LDCGSHCKGSDAHCCR-PLNTGANSGLVQRR--IQSI-----WHTCAENWTQOT 726
 QY 134 AETACRQMGY----SSKPTFAVEIGPDODLVEITENSQELMRNSSGCLSGSLVSL 189
 DB 727 SDDVQCQLGLGTGNSNMPFESSG-GP-----FKIMTAPRGSLITLTADECFPDSLLTL 780

[illegible]

RESULT 2

enteropeptidase [EC 3.4.21.9] precursor [validated] - human
 N:Alternate names: enterokinase
 C:Species: Homo sapiens (man)
 C:Date: 19-May-1995 #sequence_revision 09-Aug-1996 #text_change 09-Jul-2004
 C:Accession: A56318; B43090
 R:Klamocto, Y.; Veille, R.A.; Doms
 Biochemistry 34, 4562-4568, 1995
 A:Title: cDNA sequence and chromosomal localization of human enterokinase, the proteolytic
 A:Reference number: A56318; MUID:95234679; PMID:7718557
 A:Accession: A56318
 A:Molecule type: mRNA
 A:Residues: 1-1019 <KIT>
 A:Cross-references: UNIPROT:P98073; UNIPARC:UPI00003FE65; GB:U09860; NID:9746412; PIDN:
 Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
 A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease compo
 A:Reference number: A43090; MUID:94323561; PMID:8052624
 A:Accession: B43090
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 749-1019 <KIT>
 A:Cross-references: UNIPARC:UPI0000172B0B; GB:U09860
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 (noted below) or with amino-terminal myristoylation of the heavy chain.
 C:Genetics:
 A:Gene: GDB:PRSS7
 A:Cross-references: GDB:384083; OMIM:226200
 A:Map position: 21q21-21q21
 C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and light)
 ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms involv
 ducts.
 C:Function:
 A:Description: cleaves activation peptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolase cascade
 C:Superfamily: enteropeptidase; Clt/Cls repeat homology; LDL receptor ligand-binding rep
 C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen
 F.1-784/Product: enteropeptidase heavy chain #status predicted <HCH>
 F.122-38/Domain: transmembrane #status predicted <TM>
 F.184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F.342-504/Domain: MAM homology <MAM>
 F.526-631/Domain: Clt/Cls repeat homology <CLR>
 F.643-677/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F.678-783/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SR
 F.785-1019/Product: enteropeptidase light chain #status predicted <LCH>
 F.785-1014/Domain: trypsin homology <TRY>
 F.116, 147, 179, 328, 335, 388, 440, 470, 503, 534, 630, 682, 706, 725, 848, 887, 909, 949/Binding site:
 F.1772-896, 810-826, 910-977, 941-956, 967-995/Disulfide bonds: #status predicted
 F.1825, 876, 971/Active site: His, Asp, Ser #status predicted

```

Query Match      24.9%; Score 582.5; DB 1; Length 1019;
Best Local Similarity 33.8%; Pred. No. 56-40;
Matches 135; Conservative 62; Mismatches 132; Indels 71; Gaps 15;

QY HF-----IPKQCDGELDCPLGDEDEHCVKSPPEGPAVAVRLSKDSTLQVLDSATG 119
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 647 HQCKNGECVFLPVNLCDGHLCEDSDDEADCYRFF-----NGATTN 686

QY 120 N-----WFSACPDNFTEALAEACRWGY----SSKPTRAVEIGDDDLDAVE 164
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 687 NNGLVRFRIQSIMTHRCANMTTQISNDVCQLLGSGNSSRPIF-STDGCGFYLVNTAP 745

QY 165 ----ITENQELRMRRNSGPCLSGLSVSLHC--LAGCKSLK----TPRVGGEASVDSW 214

```

Db	746	DGHLILPSSQ-----CLDLSILTQCNHKSCKKLLAAQDITPKIVGSSNNKEGAM	796
Oy	215	PMQVSIQDKOHVCGGSLDPHWVLTAHC-FRKHTDVFNKKVRAGSDKLSGFPSLAVAK	273
Db	797	PMVVLGYGGGLTLCGASLVSDDMLVSAHCVYGRNLKESKPTALIGLHNKSULTSPQIV	856
Oy	274	II--I--I--EENPMY-----PKNDIALMKLQPLTFSGIVRI--CLPFFDEELTPATPLWITGNG	328
Db	857	RLIDEIVNPHYNRRRKNDIAMHLEKNAVYDIOEICLPEENQVPPFGNCSJAGAG	916
Oy	329	FTKNGGKMSIILIOASQVUIDSTRCNADDAVQGEVTEKMKCAGIPEGGVPTCOGDSGAP	388
Db	917	-TVVYQGTAILIQEADVPLLSNERCQ--QCMPEYVITENMICAGYBEGSIDCQDSGSP	974
Oy	389	LMYQ-SDQMAYVGVSMGTCGGSPSTPBYVTVSAVLMVI	427
Db	975	LMQENNRWFLAGVTSFGYKCALPRRPGYAVAVSFFTEMI	1014

RESULT 3

enteropeptidase (EC 3.4.21.9) precursor [validated] - bovine
 N:Alternate names: enterokinase
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: A43090; A48874; A61436
 R:Kikamori, Y.; Yuan, X.; Wu, O.; McCourt, D.W.; Sadler, J.E.
 Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994
 A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of two distinct domains
 A:Reference number: A43090; MUID:94329561; PMID:8052624
 A:Accession: A43090
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1035 <KIT>
 A:Cross-references: UNIPROT:P98072; UNIPARC:UPI00004BBB5; GB:U09859; NID:G746410; PID:N-1035
 A:Experimental source: small intestine
 R:Valle, E.R.; Rehmtulla, A.; Racie, L.A.; DiBlasio, E.A.; Ferenz, C.; Grant, K.L.; et al.
 J. Biol. Chem. 268, 23311-23317, 1993
 A:Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of bovine enterokinase
 A:Reference number: A48874; MUID:94043122; PMID:8226655
 A:Accession: A48874
 A:Molecule type: mRNA
 A:Residues: 801-1035 <LAV>
 A:Cross-references: UNIPARC:UPI00001113D; GB:U19663; NID:9416131; PID:N-AAA6035.1; PID:G746410
 A>Note: parts of this sequence, including the amino end of the mature protein, were confirmed by sequencing
 U. J. Protein Chem. 10, 475-480, 1991
 A:Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.
 A:Reference number: A61436; MUID:92189715; PMID:1799406
 A:Accession: A61436
 A:Molecule type: protein
 A:Residues: 801-807, 'Y', 809-827 <LIG>
 A:Cross-references: UNIPARC:UPI00001468A0
 C:Comment: The mechanism of association with the membrane of the intestinal brush border
 membrane attachment using a signal-anchor sequence.
 C:Comment: Conversion from membrane-bound to soluble forms may involve further processing.
 C:Complex: mature enteropeptidase is variously reported to contain two (heavy and light) subunits
 C:Function: 1
 A:Description: cleaves propeptide from trypsinogen to produce active trypsin
 A:Pathway: intestinal digestive hydrolase cascade
 C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding repeat
 C:Keywords: glycoprotein, hydrolase, predicted, serine proteinase, transmembrane protein
 F:122-38/Domain: transmembrane #status predicted <TM>
 F:122-117/Product: enteropeptidase mini chain #status predicted <MC>
 F:118-800/Product: enteropeptidase heavy chain #status predicted <HC>
 F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>
 F:358-520/Domain: MAM homology <MAM>
 F:542-647/Domain: C1r/C1s repeat homology <C1R>
 F:559-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>
 F:694-799/Domain: scavenger receptor cysteine-rich domain homology <SRC>
 F:801-1035/Product: enteropeptidase light chain #status predicted <LC>

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:42:54; Search time 39 seconds

(without alignments)
1073.187 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPSSDQPLNSLDVKKPKRP.....VYTKVSAYLNMVYWKAEI 435

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	584.5	25.0	417	1	S00845	hepsin (EC 3.4.21.
2	582.5	24.9	1019	1	A56318	enteropeptidase (E
3	571	24.4	1035	1	A43090	hepsin (EC 3.4.21.
4	568.5	24.3	416	1	S33777	hepsin (EC 3.4.21.
5	552	23.6	1034	1	A53663	enteropeptidase (E
6	511	21.8	638	1	K0RTP1	plasma kallikrein
7	508	21.7	638	1	K0MSPL	plasma kallikrein
8	505	21.6	436	2	U00172	polyprotein - Atr1
9	504.5	21.5	1524	2	T30337	acrosin (EC 3.4.21
10	503	21.5	418	2	A37404	acrosin (EC 3.4.21
11	501.5	21.4	437	2	S18407	acrosin (EC 3.4.21
12	493.5	21.1	421	2	S29599	acrosin (EC 3.4.21
13	491.5	21.0	638	1	K0RTP1	plasma kallikrein
14	488.5	20.9	1113	2	U00172	polyprotein - Atr1
15	481	20.5	421	1	S11674	acrosin (EC 3.4.21
16	478.5	20.4	415	1	A31170	acrosin (EC 3.4.21
17	474	20.2	855	2	U00172	polyprotein - Atr1
18	470	20.1	761	2	U00172	polyprotein - Atr1
19	468.5	20.0	431	2	S47538	acrosin (EC 3.4.21
20	461.5	19.7	277	2	S35340	chymotrypsin (EC 3
21	460.5	19.7	263	1	KYRTB	chymotrypsin (EC 3
22	458.5	19.6	267	2	S40006	trypsin (EC 3.4.21
23	458.5	19.6	267	2	S40006	trypsin (EC 3.4.21
24	457.5	19.5	625	1	K0RTP1	plasma (EC 3.4.21
25	455.5	19.4	420	2	A55283	acrosin (EC 3.4.21
26	449	19.2	420	2	A55283	acrosin (EC 3.4.21
27	448	19.1	810	1	A57014	proteasin (EC 3.4.
28	447.5	19.1	343	1	A57014	proteasin (EC 3.4.
29	445.5	19.0	455	2	A61545	plasmin (EC 3.4.21

ALIGNMENTS

30	445	19.0	263	2	A31299	chymotrypsin (EC 3
31	443	18.9	274	2	S35339	trypsin (EC 3.4.21
32	443	18.9	275	2	S40007	trypsin (EC 3.4.21
33	441.5	18.9	263	2	A21195	chymotrypsin (EC 3
34	439.5	18.8	810	1	PLHU	plasmin (EC 3.4.21
35	438	18.7	812	1	PLHO	plasmin (EC 3.4.21
36	432	18.4	245	1	KYBOA	chymotrypsin (EC 3
37	431.5	18.4	810	2	B30848	plasmin (EC 3.4.21
38	430	18.4	275	2	S40005	trypsin (EC 3.4.21
39	427.5	18.3	460	2	B61545	plasmin (EC 3.4.21
40	426	18.2	269	2	B26823	plasmin (EC 3.4.21
41	425.5	18.2	2616	2	A57096	pancreatic elastase
42	424	18.1	238	1	TRWV5Y	trypsin-like prote
43	424	18.1	271	1	EMRT2	pancreatic elastase
44	422	18.0	4548	1	S00657	apoptin (a) (EC
45	419.5	17.9	250	2	S55493	serine proteinase

RESULT 1
S00845
hepsin (EC 3.4.21.-) - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004
C/Accession: S00845
Rileyus, S.P.; Loeb, K.R.; Hagen, F.S.; Kurachi, K.; Davie, E.W.
Biochemistry 27, 1067-1074, 1988
A/Title: A novel trypsin-like serine protease (hepsin) with a putative transmembrane do
A/Reference number: S00845; MUID:88209431; PMID:2835076
A/Accession: S00845
A/Molecule type: mRNA
A/Residues: 1-417 <LET>
A/Cross-references: UNIPROT:P05981; UNIPARC:UPI00003FE67; EMBL:X07732; NID:G32063; PIDN
A/Genes: GDB:HPN; TMPSRSL; hepsin
A/Map position: 19q11-19q13.2
C/Keywords: hydrolase; liver; serine proteinase; transmembrane protein
F/23-45/Domain: transmembrane #status predicted <TMN>
F/163-400/Domain: trypsin homology <TRY>
F/188-204,291-359,322-338,349-381/Disulfide bonds: #status predicted
F/203,257,353/Active site: His, Asp, Ser #status predicted

Query Match 25.0%; Score 584.5; DB 1; Length 417;
Beet Local Similarity 31.5%; Pred. No. 1.2e-40;
Matches 140; Conservative 60; Mismatches 161; Indels 83; Gaps 11;

QY	19	KPRIPMETFRKVGPIIILALSLAIIIVVLIKVLIDKTYFLCGQPLHPIPRKQLCDGE	78
DB	15	RPKVAALT---AGTLTLTRIGASWVAIVLIR-----	45
QY	79	LDCEPLGEDEHCVKSPFEGPAVAVRLSKDRSTLOVLSATGMWFSACPDNTEALATAC	138
DB	46	-----SDGE-----PLYPVQVSSADARLMVDTKEGTWRLICSSRSNARVAGLSC	90
QY	139	RQMGYSKPTFRFAVEIIGPDQDDVVEITEN-----SQELRMNNSGPGCLTS	183
DB	91	EEWGF-----LRAL---TSELDVTRTAGANGTSGFPCVDEGRLLPHQRLLEVISVDCDR	142
QY	184	GSLVSLHCLACG-KSLKTPRVVGGEEASVDSMPQVSYIDYQAHVCGGSLIDPHAVLTAA	242
DB	143	GRFLALICDDCGRRKLPVDRIVGGRVTSIGRWPMQVSLYDGAHLCGGSLASGDVLTAA	202
QY	243	HCFRKATDVFN-WKVYAGSDKLGSPFSLAVAKIIIEFNPMVP-----KDNDAIMLK	293
DB	203	HCFPERNRVLSRMRYVAGVAAQASPHGLQVIGQAVVYHCGVLPFPDDPNSENSENDAIVH	262
QY	294	LQPLTFSGTVPICLPFDEBELTPATPLMILIGMFTKQNGSKMSDIILOASVQVYDSTR	353
DB	263	LSPPLPTEVIOVPCVPAAGALVDGKICTVIGWNT-QYVGQQAQGVIGEARVPIISNDV	321

THIS PAGE BLANK (USPTO)

Qy 405 GYCGGPGSTPGVYTKVSAIYLMWYVWK 432
Db 230 GIGCARPKPKPGVYVRSRFTWIEGVNR 257

RESULT 12

US-11-056-621-4
; Sequence 4, Application US/11056621
; Publication No. US20050262592A1
; GENERAL INFORMATION:
; APPLICANT: Spencer, David
; APPLICANT: Dickey, Lynn F.
; APPLICANT: Gaadaska, John R.
; APPLICANT: Wang, Xiaowei
; APPLICANT: Cox, Kevin M.
; APPLICANT: Peele, Charles G.
; TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
; TITLE OF INVENTION: MICROPLASMINOGEN IN DUCKWEED
; FILE REFERENCE: 40989/274646
; CURRENT APPLICATION NUMBER: US/11/056,621
; PRIOR FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: US 60/543,487
; PRIOR FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of mature human plasminogen
US-11-056-621-4

Query Match 18.7%; Score 438.5; DB 7; Length 791;

Best Local Similarity 37.6%; Pred. No. 2.3e-36;
Matches 94; Conservative 43; Mismatches 94; Indels 19; Gaps 6;

Qy 194 CGKSLKTP----RYVGGEEASVDSWPMQVSIQVD-KQHVCGGSLDPHMTLTAHCRK 247
Db 548 CGKQVBEKPKCPGRVGCVAHPHSPWQVSLRTFGHFCGGTILSPHWLTAHCLERK 607
Qy 248 HTDVFNMKVRAGSDLSF----PSLAVAKIIIEFNPMYKNDIAIMKIQPLTFSGT 303
Db 608 SPRBSYVVLGAHQEVLEPHVOIEVSRLE----LEPRKDIALLKLSPPAVITDK 661
Qy 304 VRPICLPFDEBELPATPLMTIIGWFTKONGKMSDILLQASVOYIDSTRGNADAYOGS 363
Db 662 VIPACLPSPNVVADRTCEFTVGGETGTG--AGLKEADLPVTEKVCNRYEFLNGR 719
Qy 364 VTEKMMCAIGPEGVDTCQDSDGGPLM-YQSDQMHVGVISWGYCGGSPSTPGVYTKVSA 422
Db 720 VQSTELCAGHLAAGTDSQDSDGGPLVCFEKKYILQGVTSGLGACARPNKPGVYVRSR 779
Qy 423 YLMWYVWK 432
Db 780 FVTWIEGVNR 789

RESULT 13

US-10-131-826A-506
; Sequence 506, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Berezhni, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 506
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-131-826A-506

Query Match 15.9%; Score 371.5; DB 6; Length 250;

Best Local Similarity 35.2%; Pred. No. 2.5e-30;
Matches 90; Conservative 40; Mismatches 99; Indels 27; Gaps 9;

Qy 189 LHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQVDKQHVCGGSLDPHMTLTAHCRK- 247
Db 8 LLAATAGVGETETIINGFECKPHSQPQALFEKTRLLCGATLILAPRWLLTAHCLKPR 67
Qy 248 ---HTDVFNMKVRAGSDK---LGSFSLAVAKIIIEFNPMYK-D--NDIAIMKIQP 297
Db 68 YIVHLGQHNLQKEGCGQTRATESFPHPG-----FNNSLPNKHNDIMLVKMASP 119
Qy 298 LTFSGYVRPICLPFDEBELPATPLMTIIGWFTKONGKMSDILLQASVOYIDSTRGNAD 357
Db 120 VSIWAVRPLTL--SRCVTAGTSCLLISGWSTSSPOLRLPHTLRCAVITLIEHOKC--E 175
Qy 358 DAYQGEVTEKMMCAIGPEGVDTCQDSDGGPLMTYQSDQMHVGVISWGYG-CGSPSTPGV 416
Db 176 NAYPQNTIDTMVCAISVQEGGDSQDSDGGPLVNGS---LQSIISMGDCAITRKRGV 232
Qy 417 YTKVSAIYLMWYVWK 432
Db 233 YTKVCKVVDWIQETMK 248

RESULT 14

US-11-029-003-6
; Sequence 6, Application US/11029003
; Publication No. US20050260194A1
; GENERAL INFORMATION:
; APPLICANT: PETERS, ROBERT T.
; APPLICANT: MEZO, ADAM R.
; APPLICANT: RIVERA, DANIEL S.

```
GENERAL INFORMATION:
APPLICANT: Ma, Jian-xing
TITLE OF INVENTION: COMPOUNDS USEFUL IN INHIBITING VASCULAR LEAKAGE, INFLAMMATION AND
FILE REFERENCE: 5820, 656
CURRENT APPLICATION NUMBER: US/11/010,874
PRIOR FILING DATE: 2004-12-13
PRIOR APPLICATION NUMBER: 10/963,115
PRIOR FILING DATE: 2004-10-12
PRIOR APPLICATION NUMBER: 60/510,620
PRIOR FILING DATE: 2003-10-10
PRIOR APPLICATION NUMBER: 60/528,647
PRIOR FILING DATE: 2003-12-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.3
SEQ ID NO 1
LENGTH: 812
TYPE: PRT
ORGANISM: murine
US-11-010-874-1
```

```
Query Match          19.4%; Score 455.5; DB 7; Length 812;
Best Local Similarity 39.4%; Pred. No. 4.6e-38;
Matches 102; Conservative 43; Mismatches 89; Indels 25; Gaps 8;
```

```
QY 181 CUGSLVSLHCLACGKSLKTP-----RVVGGEASVDSWPMQVSI--QYDKOHVCGSII 233
DB 560 CASAS-----SFEQGRQVEPKKCPGRVVGCVANPHSWFMQSLKTRFGQIFCGSTLI 614
QY 234 DPHWVLTAAHCFKRDVFNWKKRAGSDK--IG-SFPLSAVAKIIIEFNPMYPRKNDI 289
DB 615 APFWVLTAAHCLKRSRPFYKYLGAHEVYIRGLDQVETSAKLI-----LEPNRRDI 668
QY 290 ALMKLOPFLFSGTVPRICLPFDEBELTPATPLMITIGWETKONGKMSDILIQASVOYI 349
DB 669 ALMKLSRPAITTTKVPACLPSPNVYVADRTICYITIGWETQGTFG--AGRLKEAQLPVI 726
QY 350 DSTRCNADAVYQGEVTEKMKACGIPREGVDTCQDGGSGPLM--YQSDOMHVGVISWYGC 408
DB 727 ENKVCNRVEYIANNRVSTELCAQGLAGVDSCGDGGPLVCPEKXKYLIGVTSWGLCC 786
QY 409 GGPSTGGVTVKYSAYLNTI 427
DB 787 ARPNKGVYVRSRFDWI 805
```

```
RESULT 10
US-10-821-234-1043
Sequence 1043, Application US/10821234
Publication No. US20050255114A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmant, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
PRIOR FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PL SEQ_genes Version 1.0
SEQ ID NO 1043
LENGTH: 393
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1043
```

```
Query Match          19.2%; Score 449; DB 6; Length 393;
Best Local Similarity 36.6%; Pred. No. 7.6e-38;
Matches 113; Conservative 40; Mismatches 116; Indels 40; Gaps 10;
```

```
QY 147 PTFRAVEI--GPDQDLVVEITENSQELMRNMSGPCLSGLVSLHCLA----- 193
DB 35 PTLRAASLQGP-----LSWANAQKVILGGQIGAAVILLYGLRSGGAEGA 83
QY 194 ---CGKSLTPRVVGGEASVDSWPMQVSIQYDKOHVCGSIIIDPHWVLTAAHCFKRD 250
DB 84 EAPCGVAPQA-RITGSSSAVAGQWPMQVSIIEGVHVGGSIVSQWVLTAAHCFKRD 142
QY 251 VFNWKKRAGSDKLGSEFPLSA-VAKIIIEFNPMYPRD--NDIAMLKLOPFLFSGTVPR 306
DB 143 KEAYEVKGAHQLDSEBDAKYSTIKDIIHPSPYLQESQGDIALQLSRPTTFERYIRP 202
QY 307 ICLPFDEBELTPATPLMITIGWETKONGKMS-DILQASVOYIDSTRCMA---DDAQG 362
DB 203 ICLPAAASFPNGHCTVYTGWGHVAPSVSLTPKPLQDEVPFLISRETGNCLYNIDAKE 262
QY 363 E---VTEKMKACGIPREGVDTCQDGGSGPLMYQSD--QMHVVGIVSMGYGCGSPSTPGYTT 418
DB 263 EPHFVQEDWVCAGVYEGGKDAQGDGSGPLSCPVGGLMYLIGVISWGDACARRNPVYTT 322
QY 419 KVSAYLNTI 427
DB 323 IASSYASWT 331
```

```
RESULT 11
US-11-056-621-6
Sequence 6, Application US/11056621
Publication No. US20050262592A1
GENERAL INFORMATION:
APPLICANT: Spencer, David
APPLICANT: Dickey, Lynn F.
APPLICANT: Gasdaska, John R.
APPLICANT: Cox, Xiaowei
APPLICANT: Wang, Kevin M.
APPLICANT: Peelle, Charles G.
TITLE OF INVENTION: EXPRESSION OF PLASMINOGEN AND
FILE REFERENCE: 40989/274646
CURRENT APPLICATION NUMBER: US/11/056,621
PRIOR FILING DATE: 2005-02-11
PRIOR APPLICATION NUMBER: US 60/543,487
PRIOR FILING DATE: 2004-02-11
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 259
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Sequence of mature human microplasinogen
US-11-056-621-6
```

```
Query Match          18.9%; Score 442.5; DB 7; Length 259;
Best Local Similarity 36.2%; Pred. No. 1.9e-37;
Matches 97; Conservative 44; Mismatches 98; Indels 29; Gaps 7;
```

```
QY 176 NSSGPCLSGLVSLHCLACGKSLKTP-----RVVGGEASVDSWPMQVSIQYDKOHVCG 229
DB 8 NTQGPSPD-----CGKQVEPKKCPGRVVGCVANPHSWFMQSLKTRFGQIFCG 57
QY 230 GSIIIDPHWVLTAAHCFKRDVFNWKKRAGSDKLSF---PSIAVAKIIIEFNPMYPRK 285
DB 58 GTLISPFWVLTAAHCLKRSRPFYKYLGAHGEVNLPEHVQIEVERLP-----LEPT 111
QY 286 DNDIAMLKLOPFLFSGTVPRICLPFDEBELTPATPLMITIGWETKONGKMSDILIQAS 345
DB 112 RKDIALKLSPPAVITTKVPACLPSPNVYVADRTICYITIGWETQGTFG--AGRLKEAQ 169
QY 346 VQVIDSTRCNADAVYQGEVTEKMKACGIPREGVDTCQDGGSGPLM--YQSDOMHVGVISW 404
DB 170 LPVIENTCNRYEFLNKRVOSTELCAQGLAGVDSCGDGGPLVCPEKXKYLIGVTSW 229
```


Db 413 LRDM 417

RESULT 7

US-11-137-465-56

; Sequence 56, Application US/11137465
; Publication No. US20050255558A1

; GENERAL INFORMATION:

; APPLICANT: Agarwal, Pankaj
; APPLICANT: Murdock, Paul R.
; APPLICANT: Rizvi, Safia K.
; APPLICANT: Smith, Randall, F.
; APPLICANT: Xiang, Zhaoying
; APPLICANT: Kadnick, Karen

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GPs0018

; CURRENT APPLICATION NUMBER: US/11/137,465

; CURRENT FILING DATE: 2005-05-25

; PRIOR APPLICATION NUMBER: US/10/239,663

; PRIOR FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: PCT/US01/09226

; PRIOR FILING DATE: 2001-03-22

; PRIOR APPLICATION NUMBER: 60/192,158

; PRIOR FILING DATE: 2000-03-24

; PRIOR APPLICATION NUMBER: 60/192,668

; PRIOR FILING DATE: 2000-03-27

; PRIOR APPLICATION NUMBER: 60/200,166

; PRIOR FILING DATE: 2000-04-27

; NUMBER OF SEQ ID NOS: 66

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 56

; LENGTH: 305

; TYPE: PRT

; ORGANISM: Homo sapiens

US-11-137-465-56

Query Match 20.2%; Score 473.5; DB 7; Length 305;

Best Local Similarity 37.8%; Pred. No. 1.8e-40;

Matches 95; Conservative 43; Mismatches 100; Indels 13; Gaps 4;

QY 193 ACGSLKTPRVVGEASVDSMPQVSIQYDKOHVCGSILDPHVLTAACFRHTDVF 252

Db 32 ACGREIHALVAGVGSARGRPWQASLRRLRRHCGSLSRRWVLSAHCFOGHYPS 91

QY 253 NMKTRAGS-----DKLSFSPSLAVAKIIIEFNPMYPKNDIALMLQPLTFSGTVR 305

Db 92 EMTYQLGSLTRPFPWNLRAVSSRKVQDIIIVNPALGLVANDIALRLASSVTYNATIQ 151

QY 306 PICLPFDEBLTPATPLMIIGWFTKONGKMSD--ILLQASVOYIDSTRCN--ADDA 360

Db 152 PICIESSTFNVRHDCVVTGWLISPGTPLRPYYNLREAQVTLNTRCNVLEEGPSS 211

QY 361 QGEVTEKMKAGIPBGGVDTCCGDSGGPLMTQSD-QMHVVGIVSGKCGGSPSTPGVYTK 419

Db 212 RSMWDSWFCGAEBSVDTCKDSCGSLVCDKGLMYQVIGVSGMDCCGPNRPGVYTN 271

QY 420 VSAVLMWLYN 430

Db 272 ISVYFHWIRRV 282

RESULT 8

US-10-131-826A-222

; Sequence 222, Application US/10131826A

; Publication No. US20050245730A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Demoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C128

; CURRENT APPLICATION NUMBER: US/10/131,826A

; CURRENT FILING DATE: 2002-04-24

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059117

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059122

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059184

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059263

; PRIOR FILING DATE: 1997-09-18

; PRIOR APPLICATION NUMBER: 60/059352

; PRIOR FILING DATE: 1997-09-19

; PRIOR APPLICATION NUMBER: 60/059388

; PRIOR FILING DATE: 1997-09-19

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 222

; LENGTH: 290

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-131-826A-222

Query Match 20.0%; Score 467.5; DB 6; Length 290;

Best Local Similarity 37.1%; Pred. No. 6.9e-40;

Matches 99; Conservative 49; Mismatches 92; Indels 27; Gaps 9;

QY 184 GSLVSLHCLACGSLKTPRVVGEASVDSNPQVSIQYDKOHVCGSILDPHVLTPAH 243

Db 16 GSRKAKATACGRBRMLNRMVWGQDTQEGEPWQVSIQRNGSHFCGSLIAEQVLTAAH 75

QY 244 CFRKHTDVFNKVAAGSDKL-GSFPSLAVAKIIIEFNPMY---PKNDIALMLQPLT 299

Db 76 CFRNTSSTLYQVLLGLRQLQPPHMYARVRQVENSPLQGTASSADVALVLEAVP 135

QY 300 FSGTVRPLCLPFPBELTPATPL--WIIIGWFTKONGKMSD-----ILLQASVOYIDS 351

Db 136 FTNYILPVCUP--DPSVIFETGMNCWVGWSPSE-----EDLLPPEITLQKLAVPITDT 188

QY 352 TRCN-----ADDAVQGE-VTEKMKAGIPBGGVDTCCGDSGGPLMTQSDQ-WHVVGIVS 403

Db 189 PKCNLAYSXDEPFGYQPKTIKNDLACGFEKGGKDACGDSGSLVCLVQGSWILQAGVIS 248

QY 404 WGYCGGSPSTPGVYTKYSAYLMWLYN 430

Db 249 WEGECARQNRPGYIVRVTAHNMWTHRI 275

RESULT 9

US-11-010-874-1

; Sequence 1, Application US/11010874

; Publication No. US20050250694A1

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 4
; LENGTH: 1113
; TYPE: PRT
; ORGANISM: Mus Musculus
US-11-067-811-4

Query Match 20.9%; Score 488.5; DB 7; Length 1113;
Best Local Similarity 33.2%; Pred. No. 3.4e-41;
Matches 130; Conservative 52; Mismatches 148; Indels 61; Gaps 17;

QY 69 IPRKOLCDGELDCLPGEDEEHCVKSPBEGPAVAVALRSKDRSTLQVLDATGMMFSAQCFDN 128
DB 736 VPRDLMCDGWDVCSDDSCSDMGCVTLSSKNG-----NSSSLTLTVKSAKEH--HVCADG 785
QY 129 FTEALAEATACROMGYSSKRTFRAVEIGPRODDLVTEITNSQELRN---RMSG----- 179
DB 766 WRETLSQLCKQMGJGEPSTVKLI--PGQ-----BGQOWLRLYPMWENLNGSTLOE 834
QY 180 -----PCLSGSLVSLHCLA--CGK---SLKTPRVVGESEASVDSMPQVSIQYDKQ-HV 227
DB 835 LLVYHSCGSRSEISLCKQDCGRPARAAMKRILIGRTSRGRMPQCSLQSEPSGHI 894
QY 228 CGGSLDPHMLVLAHCERKHTDVFNMKYRAGSDKL--GSPFSLAVAKIIIEFNPMYP 284
DB 895 CGCVLIARKKVLTVAHCFEGREDADVMKVFGINNLDHPSGFQTRFVKTILL--HPRYS 952
QY 285 K---DNDIALMKLOPFLTSGTVRPICLPFDEBLTPATPLMIGMFTKONGKAKSDIL 341
DB 953 RAVVYDIDSVBELSDINETSYYRVPCLSPPEYLEPDDYCYITGNG--HMGNKKPFKL 1009
QY 342 LQASVOVISTRCNADAY--QGEVTEKMKACAGIPREGVDTCCGDSGGPLMYQ--SDQW 397
DB 1010 QGEVRIITPLEQC--QSYFDMKTIITRMICAGYESGTVDSMGDSGGPLVCERPGQNT 1066
QY 398 VVGIVSMGYGCGGPGST-PGVYTKVSAYLMI 427
DB 1067 LFGLTSMGSCVPSKVLGPGVYSNVSFYEWMI 1097

RESULT 5
US-11-067-811-1
; Sequence 1, Application US/11067811
; Publication No. US20050260688A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; APPLICANT: Enshell-SciJiffers, David
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; PRIOR FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; PRIOR FILING DATE: 2004-02-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1
; LENGTH: 1042
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-067-811-1

Query Match 20.7%; Score 484; DB 7; Length 1042;
Best Local Similarity 32.6%; Pred. No. 8.9e-41;
Matches 126; Conservative 64; Mismatches 146; Indels 50; Gaps 17;

QY 69 IPRKOLCDGELDCLPGEDEEHCVKSPBEGPAVAVALRSKDRSTLQVLDATGMMFSAQCFDN 128
DB 666 VPRDLMCDGWDVCSDDSCSDMGCVTLSSKNG-----TISINNS-SFLMHRATEH--HVCADG 717
QY 129 FTEALAEATACROMGYSSKRTFRAVEIGPRODDLVTEITNSQELRN---RMSG----- 178
DB 718 WRETLSQLCKQMGJGEPSTVKLIQ---DQEKPRMLTLHSWNESLNGTTLHELVLVNGS 774

QY 179 GPCLSGLVSLHCLA--CGK---SLKTPRVVGESEASVDSMPQVSIQYDKQ-HVCGGSI 232
DB 775 --CESRSKISLILCTKQDCGRPARAAMKRILIGRTSRGRMPQCSLQSEPSGHI 832
QY 233 LDPHMLVLAHCERKHTDVFNMKYRAGSDKL--GSPFSLAVAKIIIEFNPMYPK--D 286
DB 833 IAKKVVLTVAHCFEGRENAAVKVLGINNLDHPSVFMQTRFVKTIIL--HPRYSRAVD 890
QY 287 NDIALMKLOPFLTSGTVRPICLPFDEBLTPATPLMIGMFTKONGKAKSDILQASV 346
DB 891 YDISIVESEDISESGVYRVPCLPPEOMLEBDYCYITGNG--HMGNKKPFKQEGEV 947
QY 347 QVISTRCNADAY--QGEVTEKMKACAGIPREGVDTCCGDSGGPLMYQ--SDQW 402
DB 948 RILSLHLC---QSYFDMKTIITRMICAGYESGTVDSMGDSGGPLVCERPGQNT 1004
QY 403 SMGYCGGPGST-PGVYTKVSAYLMI 427
DB 1005 SMGSCVPSKVLGPGVYSNVSFYEWMI 1030

RESULT 6
US-11-102-240-106
; Sequence 106, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; PRIOR FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 106
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-106

Query Match 20.3%; Score 475; DB 7; Length 423;
Best Local Similarity 39.6%; Pred. No. 2.1e-40;
Matches 97; Conservative 41; Mismatches 97; Indels 10; Gaps 5;

QY 190 HCLAG----KSL-KTPRVVGESEASVDSMPQVSIQYDKQHVCGGSLDPHMLVLAHC 244
DB 176 HC--CGTRRSKTLGSLRIIVGSTEVEEGEEMPOQASLDWDSHRCATLINAATWLVAHLC 233
QY 245 FRKHTDVFNMKYRAGSDKLSPFSLAVAKIIIEFNPMYPKNDIALMKLOPFLTSGTV 304
DB 234 FTYTKNPARMTASFEVITIKPSMKRGLRIIVHEKXKIPSHDYDLSLAEISSPYVTNAV 293
QY 305 RPICLPFDEBLTPATPLMIGMFTKONGKAKSDILQASVOVISTRCNADAYQGEV 364
DB 294 HRVCLPDASVEYEPQGDVNFVTFGALK-NDGYSQWHLQAQVTLIDATTCNEPQAYNDAL 352
QY 365 TEKMKACAGIPREGVDTCCGDSGGPLMYQ--DQWTVGIVSMGYGCGGPGSTPGVYTKYSA 422
DB 353 TPRMLCAASLLEKTDACOGDSGGPLVSSDARDIWTLAGIVSGDECAKPNKPGVYTRVYA 412
QY 423 YLNM 427

Db 179 LSGSLVSLHCLACGSKLTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTA 238
Qy 242 AHCFRKHDTVDENKVRGSDKLSFSPSLAVAKIIIEFNMPYKNDIAMLKQPLTFPS 301
Db 239 AHCFRKHDTVDENKVRGSDKLSFSPSLAVAKIIIEFNMPYKNDIAMLKQPLTFPS 298
Qy 302 GTVRPCLPFDEELTPATPLMIIGMFTKQNGKMSDILLQASVQYIDSTRCNADAYQ 361
Db 299 GTVRPCLPFDEELTPATPLMIIGMFTKQNGKMSDILLQASVQYIDSTRCNADAYQ 358
Qy 362 GEVTEKMKACGIPREGVDTCQGDGSGPLMTQSDQMHVGVISWGYCGGPGSTPGVTKVS 421
Db 359 GEVTEKMKACGIPREGVDTCQGDGSGPLMTQSDQMHVGVISWGYCGGPGSTPGVTKVS 418
Qy 422 AYLNMIVNWKAKEL 435
Db 419 AYLNMIVNWKAKEL 432

RESULT 2
US-11-182-752-4
; Sequence 4, Application US/11182752
; Publication No. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 417
; TYPE: PRT
; ORGANISM: HUMAN
US-11-182-752-4

Query Match 25.1%; Score 588.5; DB 7; Length 417;
Best Local Similarity 31.5%; Pred. No. 7.7e-52;
Matches 140; Conservative 60; Mismatches 161; Indels 83; Gaps 11;

Qy 19 KPRIPMEFRKVKGPITIIALLSLASITIVVVLKYLIDKYYFLCGQPLHFIPRQQLDGE 78
Db 15 RPKVAALT---AGTLILLTAIGASWAIIVALLR----- 45
Qy 79 LDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTLOVLDATGNWFSACFDNFTALAEATAC 138
Db 46 -----SDQE-----PLYVQVSSADARLMVFDKTEGTWRLCSSRSNARVAGLSC 90
Qy 139 RQMGYSKPTFRFAVEIGPDOLDVVEITEN-----SOELRMKNSGPGCLS 183
Db 91 EEMGF-----LRAL---THSELVDVATAGNGSGFVCDGRLPHTQGLLEVVISDQPR 142
Qy 184 GSLVSLHCLACG-KSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTA 242
Db 143 GRFLAALICODCGRRKLPVDRIVIGGRDTSIGRWPMQVSLRDGAHLGCSLLSGDWLTPA 202
Qy 243 HCFRKHDTVDEN-KWYRAGSDKLSFSPSLAVAKIIIEFNMPY-----KNDIALMK 293
Db 203 HCFRKHDTVDENKVRGSDKLSFSPSLAVAKIIIEFNMPY-----KNDIALMK 262
Qy 294 LQPLTFSGTVAPICLPFDEELTPATPLMIIGMFTKQNGKMSDILLQASVQYIDSTR 353
Db 263 LSSPLRLEYIQPVCLPAAGQALVVGKICTVYGWNT-QYVGQAGVLEARVPIISNDV 321
Qy 354 CNADAYQGEVTEKMKACGIPREGVDTCQGDGSGPLMTQSDQMHVGVISWGYCGGPGSTPGVTKVS 421
Db 322 CNGAFYVGNQIKRPMKFCAGYPEGGIDACQGDGSGPVCDSISRTPRWRLCGIVSWGTGC 381
Qy 409 GGSTPGVTVKVSAYLNMIVNWK 432

Db 382 ALAQKPGVYTKVSDFREMIPOAIK 405

RESULT 3
US-11-182-752-2
; Sequence 2, Application US/11182752
; Publication No. US20050250154A1
; GENERAL INFORMATION:
; APPLICANT: Gan, Weiniu et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001194DIVII
; CURRENT APPLICATION NUMBER: US/11/182,752
; CURRENT FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 376
; TYPE: PRT
; ORGANISM: HUMAN
US-11-182-752-2

Query Match 24.8%; Score 580; DB 7; Length 376;
Best Local Similarity 30.8%; Pred. No. 4.8e-51;
Matches 132; Conservative 57; Mismatches 146; Indels 94; Gaps 9;

Qy 19 KPRIPMEFRKVKGPITIIALLSLASITIVVVLKYLIDKYYFLCGQPLHFIPRQQLDGE 78
Db 15 RPKVAALT---AGTLILLTAIGASWAIIVALLR----- 45
Qy 79 LDCPLGDEEHCVKSPFEGPAVAVRLSKDRSTLOVLDATGNWFSACFDNFTALAEATAC 138
Db 46 -----SDQE-----PLYVQVSSADARLMVFDKTEGTWRLCSSRSNARVAGLSC 90
Qy 139 RQMGYSKPTFRFAVEIGPDOLDVVEITENSQELRMKNSGPGCLSGLVSLHCLACG-KS 197
Db 91 EEMGF-----LRAL---THSELVDVATAGNGSGFVCDGRLPHTQGLLEVVISDQPR 142
Qy 198 LKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGSILDPHWLTAHCFRKHDTVDEN-KWY 256
Db 117 LPVDRIVIGGRDTSIGRWPMQVSLRDGAHLGCSLLSGDWLTPAHCPEENRRLSLMRV 176
Qy 257 RAGSDKLSFSPSLAVAKIIIEFNMPY-----KNDIALMKLQPLTFSGTVAPIC 308
Db 177 FAGVAQASPHGLQVQAVVYHGGYLPFRDPNSENSNDIALVHLSPLRLEYIQPV 236
Qy 309 LPFDEELTPATPLMIIGMFTKQNGKMSDILLQASVQYIDSTRCNADAYQGEVTEKMK 368
Db 237 LPAAGQALVVGKICTVYGWNT-QYVGQAGVLEARVPIISNDVCGADFYGNQIKPM 295
Qy 369 MCAGIPREGVDTCQGDGSGPLMTQSDQMHVGVISWGYCGGPGSTPGVTVKVSAY 423
Db 296 FCAGYPEGGIDACQGDGSGPVCDSISRTPRWRLCGIVSWGTGALAQKPGVYTKVSDF 355
Qy 424 LNMIVNWK 432
Db 356 REMIFQAIK 364

RESULT 4
US-11-067-811-4
; Sequence 4, Application US/11067811
; Publication No. US20050260688A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A
; APPLICANT: Enshell-SciJffers, David
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HAIR GROWTH
; FILE REFERENCE: 10287-083001
; CURRENT APPLICATION NUMBER: US/11/067,811
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US 60/548,272
; PRIOR FILING DATE: 2004-02-27

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:53:55 ; Search time 11 Seconds

(without alignments)
189.357 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342
Sequence: 1 MDPSDQPLNSLDVKPLRKP.....VYTKVSAYLNWYVWRAEL 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 26661 seqs, 4788334 residues

Total number of hits satisfying chosen parameters: 26661

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/2/pubppa/US09_NEW_PUB.pep.*
2: /cgn2_6/prodata/2/pubppa/US06_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubppa/US07_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubppa/US08_NEW_PUB.pep.*
5: /cgn2_6/prodata/2/pubppa/PCT_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubppa/US10_NEW_PUB.pep.*
7: /cgn2_6/prodata/2/pubppa/US11_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubppa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2297.5	98.1	432	7	US-11-102-240-112
2	588.5	25.1	417	7	US-11-182-752-4
3	580	24.8	376	7	US-11-182-752-2
4	488.5	20.9	1113	7	US-11-067-811-4
5	484	20.7	1042	7	US-11-067-811-1
6	475	20.3	423	7	US-11-102-240-106
7	473.5	20.2	305	7	US-11-137-465-56
8	467.5	20.0	290	6	US-10-131-826A-222
9	455.5	19.4	812	7	US-11-010-874-1
10	449	19.2	393	6	US-10-821-234-1043
11	442.5	18.9	259	7	US-11-056-621-6
12	438.5	18.7	791	7	US-11-056-621-6
13	371.5	15.9	250	6	US-10-131-826A-506
14	368	15.7	671	7	US-11-029-003-6
15	355.5	15.2	690	6	US-11-029-003-8
16	343.5	14.7	260	6	US-10-131-826A-396
17	319	13.6	236	6	US-10-821-234-1274
18	313	13.4	237	6	US-10-401-386B-24
19	307.5	13.1	238	6	US-10-821-234-1533
20	304	13.0	293	6	US-10-131-826A-456
21	297	12.9	237	6	US-10-401-386B-6
22	297	12.7	705	6	US-10-821-234-1532
23	291	12.4	255	7	US-11-147-047-36
24	285.5	12.2	194	6	US-10-401-386B-12
25	282	12.0	269	6	US-10-401-386B-2

26	231.5	9.9	301	7	US-11-147-047-37	Sequence 37, Appl
27	230.5	9.8	267	6	US-10-846-172A-2	Sequence 2, Appl
28	230.5	9.8	304	6	US-10-846-172A-5	Sequence 5, Appl
29	230.5	9.8	330	6	US-10-846-172A-7	Sequence 7, Appl
30	230.5	9.8	333	6	US-10-846-172A-6	Sequence 6, Appl
31	230	9.8	739	7	US-11-057-047-2	Sequence 2, Appl
32	230	9.8	764	7	US-11-057-047-1	Sequence 1, Appl
33	230	9.8	798	6	US-10-821-234-1034	Sequence 1034, Ap
34	229.5	9.8	255	6	US-10-846-172A-4	Sequence 4, Appl
35	226	9.6	720	7	US-11-102-240-38	Sequence 38, Appl
36	225	9.6	219	6	US-10-846-172A-1	Sequence 1, Appl
37	222	9.5	761	7	US-11-057-047-6	Sequence 6, Appl
38	217	9.3	199	6	US-10-401-386B-29	Sequence 29, Appl
39	208	8.9	241	6	US-10-131-826A-248	Sequence 248, App
40	208	8.9	243	7	US-11-137-465-38	Sequence 38, Appl
41	208	8.9	243	7	US-11-137-465-39	Sequence 39, Appl
42	194	8.3	227	6	US-10-977-334-7	Sequence 7, Appl
43	194	8.3	234	6	US-10-401-386B-8	Sequence 8, Appl
44	158	6.7	140	6	US-10-401-386B-27	Sequence 27, Appl
45	158	6.7	176	6	US-10-401-386B-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1		US-11-102-240-112	Application US/11102240
		Sequence 112, Application US/11102240	Publication No. US20050260647A1
		GENERAL INFORMATION:	
		APPLICANT: Goddard, Audrey	
		APPLICANT: Godowski, Paul J.	
		APPLICANT: Grimaldi, Christopher J.	
		APPLICANT: Gurney, Austin L.	
		APPLICANT: Wood, William I.	
		TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS	
		FILE REFERENCE: P3230R1C106C	
		CURRENT APPLICATION NUMBER: US/11/102,240	
		CURRENT FILING DATE: 2005-04-08	
		PRIOR APPLICATION NUMBER: 10/063662	
		PRIOR FILING DATE: 2002-05-07	
		PRIOR APPLICATION NUMBER: 10/006867	
		PRIOR FILING DATE: 2001-12-06	
		PRIOR APPLICATION NUMBER: PCT/US00/23328	
		PRIOR FILING DATE: 2000-08-24	
		PRIOR APPLICATION NUMBER: 60/170262	
		PRIOR FILING DATE: 199-12-09	
		NUMBER OF SEQ ID NOS: 170	
		SEQ ID NO 112	
		LENGTH: 432	
		TYPE: PRT	
		ORGANISM: Homo Sapien	
		US-11-102-240-112	
QY	Query Match	98.1%; Score 2297.5; DB 7; Length 432;	
	Best Local Similarity	98.8%; Pred. No. 1e-223; 0; Indels 5; Gaps 1;	
	Matches 429; Conservative 0; Mismatches 0;		
QY	2	MDPSDQPLNSLDVKPLRKPRIEMETFRKVGIPPIIALSLASIIIVVILIKYILDKYPL 61	
DB	4	DPDSDQPLNSLDVKPLRKPRIEMETFRKVGIPPIIALSLASIIIVVILIKYILDKYPL 63	
QY	62	CGQPLHPIRKOUCDEELDCPLGDEBEHCVKSPFGPAVAVALSKDRSTLQVLDSATGNW 121	
DB	64	CGQPLHPIRKOUCDEELDCPLGDEBEHCVKSPFGPAVAVALSKDRSTLQVLDSATGNW 123	
QY	122	FSACFDNFTALAEATACRQMGYSKPTFAVIEIGPDLDVVEITENSOELMRNMSGPC 181	
DB	124	FSACFDNFTALAEATACRQMGYS-----RAVAIGPDLDVVEITENSOELMRNMSGPC 178	
QY	182	LGSLVSLHCLACGSLKTPRVVGEBAASVDSWPMQVSIQYDKQHVCGSIIIDPMVLTJA 241	

QY	62	CGOPLHFI	PRKOLCDGBLDCPLGEBDEHCVKS	PEGPAAVAVRLSKDRSTLQVLDSATGNW	121
Db	64	CGOPLHFI	PRKOLCDGBLDCPLGEBDEHCVKS	PEGPAAVAVRLSKDRSTLQVLDSATGNW	123
QY	122	FSACFDNFT	EALAEACRQMGYSKPTFRAYEIGPDQDL	DVEITENSQELRMNRSSGPC	181
Db	124	FSACFDNFT	EALAEACRQMGYSKPTFRAYEIGPDQDL	DVEITENSQELRMNRSSGPC	183
QY	182	LSGSLVSL	HLACGKSLKTPRVGGEASVDSWPQVSI	OYDKOHVCGSILDPHVLTA	241
Db	184	LSGSLVSL	HLACGKSLKTPRVGGEASVDSWPQVSI	OYDKOHVCGSILDPHVLTA	243
QY	242	AHCFRKHTD	VNMWVRAGSDKLSFSLAVAKIIIEFNPMY	PKNDIALMKLOPPLTFS	301
Db	244	AHCFRKHTD	VNMWVRAGSDKLSFSLAVAKIIIEFNPMY	PKNDIALMKLOPPLTFS	303
QY	302	GTVRPICI	LPFDEBITPATPLMIIGWFTKONGKMSDILL	QASVQVIDSTRCNADDAVQ	361
Db	304	GTVRPICI	LPFDEBITPATPLMIIGWFTKONGKMSDILL	QASVQVIDSTRCNADDAVQ	363
QY	362	GEVTEKMW	CAGIPREGVDTCQDSCGRLMYOSDOMHVIV	ISWGYCGGSPSTPGVYTKVS	421
Db	364	GEVTEKMW	CAGIPREGVDTCQDSCGRLMYOSDOMHVIV	ISWGYCGGSPSTPGVYTKVS	423
QY	422	AYLNWIVY	VWKAEL	435	
Db	424	AYLNWIVY	VWKAEL	437	

Search completed: December 5, 2005, 13:49:12
 Job time : 192 secs

PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 XX polynucleotide that hybridizes to an ovarian cancer gene.
 PS Claim 13; Page 305; 332p; English.

CC The invention relates to a method of detecting an ovarian cancer-
 CC associated transcript in a cell from a patient, by contacting a
 CC biological sample from the patient with a polynucleotide that selectively
 CC hybridizes to a sequence at least 80% identical to any of one of 80
 CC nucleic acid sequences given in the specification. The method is useful
 CC in diagnosing ovarian cancer and in identifying and using agents and/or
 CC targets that inhibit ovarian cancer. The nucleic acid molecule,
 CC polypeptide and the antibody may also be used in detecting ovarian
 CC cancer, monitoring and early detection of relapse following treatment,
 CC monitoring response to therapy, selecting patients for post-operative
 CC chemotherapy or radiation therapy, in selecting mode of therapy,
 CC determining tumour prognosis, early detection of pre-cancerous lesions,
 CC and as vaccines. This sequence corresponds to one of the proteins used
 CC for the detection method of the invention.

CC Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 7; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-201;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSDDPINSIDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVLIDKXYFL 61
 DB 4 DPDSDDPINSIDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVLIDKXYFL 63
 QY 62 CGQPLHFTPRKQLCDGELDCPLGDEDEHCVKSPFEGAVAVRLSKDRSTLQVLDATGNW 121
 DB 64 CGQPLHFTPRKQLCDGELDCPLGDEDEHCVKSPFEGAVAVRLSKDRSTLQVLDATGNW 123
 QY 122 FSACFDNFTEALAEACROWGYSKPTFAVEIGPDODLVITNSQELMRNNSGQC 181
 DB 124 FSACFDNFTEALAEACROWGYSKPTFAVEIGPDODLVITNSQELMRNNSGQC 183
 QY 182 LSGSIVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHCGSILDPHVLTA 241
 DB 184 LSGSIVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHCGSILDPHVLTA 243
 QY 242 AHCFRKHTDVFNKVRAGSDKLSFSLAVAKIIIEFNPMYPKNDIALMLQPLTFSS 301
 DB 244 AHCFRKHTDVFNKVRAGSDKLSFSLAVAKIIIEFNPMYPKNDIALMLQPLTFSS 303
 QY 302 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAVQ 361
 DB 304 GTVRPICLPFFDEELTPATPLWIIIGWFTKONGKMSDILLQASVQVIDSTRCNADDAVQ 363
 QY 362 GEVTEKMGAGTPEGVDTCCGDSGGPLMYQSDQMHVGVISWGYGCGGPTPGVYTKVS 421
 DB 364 GEVTEKMGAGTPEGVDTCCGDSGGPLMYQSDQMHVGVISWGYGCGGPTPGVYTKVS 423
 QY 422 AYLMNIVYNWKAEL 435
 DB 424 AYLMNIVYNWKAEL 437

RESULT 15
 ADN39878
 ID ADN39878 standard; protein; 437 AA.
 AC
 XX
 XX
 ADN39878;
 DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO: C248.
 XX
 XX Human; differential expression; cancer; angiogenic disorder;
 KM fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KM inflammatory disease; autoimmune disease;

KM retinal neovascularisation syndrome; scarring; uterine fibroid;
 KM detection; diagnosis; prognosis; drug screening; drug targeting;
 KM wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KM vulnery; gene therapy; vaccine.

OS Homo sapiens.

XX NO2003042661-A2.

XX 22-MAY-2003.

PF 13-NOV-2002; 2002MO-US036810.

PR 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-034393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368099P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hervez PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;

DR WPI; 2003-468649/44.

DR N-PSDB; ADN39861.

PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosis, prognosis or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C248; 1385p; English.

CC The invention relates to nucleic acids and proteins (ADN398683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.

CC Sequence 437 AA;

Query Match 99.8%; Score 2337; DB 7; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-201;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSDDPINSIDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVLIDKXYFL 61
 DB 4 DPDSDDPINSIDVPLRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVLIDKXYFL 63

QY 362 GEVTEKMMKAGIPGCGVDTCCGDSGGPLMYOSDDQMHVVGIVSMGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMMKAGIPGCGVDTCCGDSGGPLMYOSDDQMHVVGIVSMGCGGSPSTPGVYTKVS 423
QY 422 AYLNWYINVMKAEL 435
Db 424 AYLNWYINVMKAEL 437

RESULT 13
ID ABU04915 standard; protein; 437 AA.
XX ABU04915;
AC
XX 29-JAN-2003 (first entry)
DT
XX Human expressed protein tag (EPT) #1581.
DE
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX
OS Homo sapiens.
PN MO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0328370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCO INC.
PA
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1581; 134pp; English.
PS
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 437 AA;
SQ

Query Match 99.8%; Score 2337; DB 6; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-201;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 DPDSQPLNSLDVVKLPKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVIDKTYFL 61
Db 4 DPDSQPLNSLDVVKLPKPRIPMETFRKVGIPITIIALLSLASIIIVVLKIVIDKTYFL 63
QY 62 CGGPLHFIIPRKQLCDGLDPLGDEDEHCYKSPFEGAVAVRLSKDSTQVLDASGNW 121
Db 64 CGGPLHFIIPRKQLCDGLDPLGDEDEHCYKSPFEGAVAVRLSKDSTQVLDASGNW 123
QY 122 FSACFDNFTETALATACRQMGYSKPTFRAVEIGPDODLVAETIENSQELRMNSSGPC 181
Db 124 FSACFDNFTETALATACRQMGYSKPTFRAVEIGPDODLVAETIENSQELRMNSSGPC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDKQHVCGSILDPHWVLT 241
Db 184 LSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDKQHVCGSILDPHWVLT 243
QY 242 AHCRKHTDVFNKVRAGSDKLSFSLAVAKIIIFPNPMPYPKNDIALMKLOPPLTFS 301
Db 244 AHCRKHTDVFNKVRAGSDKLSFSLAVAKIIIFPNPMPYPKNDIALMKLOPPLTFS 303
QY 302 GTVPRICLPFDELTATATPLMTIIGWGFTRKONGKMSDIILOASVOYIDRCNADAYQ 361
Db 304 GTVPRICLPFDELTATATPLMTIIGWGFTRKONGKMSDIILOASVOYIDRCNADAYQ 363
QY 362 GEVTEKMMKAGIPGCGVDTCCGDSGGPLMYOSDDQMHVVGIVSMGCGGSPSTPGVYTKVS 421
Db 364 GEVTEKMMKAGIPGCGVDTCCGDSGGPLMYOSDDQMHVVGIVSMGCGGSPSTPGVYTKVS 423
QY 422 AYLNWYINVMKAEL 435
Db 424 AYLNWYINVMKAEL 437

RESULT 14
ID ADB80525 standard; protein; 437 AA.
XX ADB80525;
AC
XX 04-DEC-2003 (first entry)
DT
XX
DE Ovarian cancer-associated protein #45.
XX
KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection.
XX
XX Homo sapiens.
OS
XX
PN WO2002102235-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (BOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC;
PI
XX WPI; 2003-167431/16.
DR
XX N-PSDB; ADB80524.
XX

CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 437 AA:

Query Match 99.8%; Score 2337; DB 6; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-201;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVYKPLKRPRIPIIIMLSTASIIIVVILKVIDKYFL 61
DB 4 DPDSQPLNSLDVYKPLKRPRIPIIIMLSTASIIIVVILKVIDKYFL 63
QY 62 CGQPLHPIPRKOLCDGLDPLGDEBEHCYKSPPEGPAVAVRUSKDRSTLQVLDATGMW 121
DB 64 CGQPLHPIPRKOLCDGLDPLGDEBEHCYKSPPEGPAVAVRUSKDRSTLQVLDATGMW 123
QY 122 FSACFDNFTALAEATACROWGYSKPTFRAVEIGPDDLDVVEITENSQELRNRSSGRC 181
DB 124 FSACFDNFTALAEATACROWGYSKPTFRAVEIGPDDLDVVEITENSQELRNRSSGRC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVVGGEASVDSWPMQVSIQYDKQHCYSGSILDPHVVLT 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVVGGEASVDSWPMQVSIQYDKQHCYSGSILDPHVVLT 243
QY 242 AHCFRKHITDVFNMKVRAAGSDKLSFSLAVAKIIIEFNPMYRKONDIALMKLQEPFLTS 301
DB 244 AHCFRKHITDVFNMKVRAAGSDKLSFSLAVAKIIIEFNPMYRKONDIALMKLQEPFLTS 303
QY 302 GTVRPICLPFDEBELTPATPLMTIIGWFTKONGKNSDILLQASVQVIDSTRCNADDAVQ 361
DB 304 GTVRPICLPFDEBELTPATPLMTIIGWFTKONGKNSDILLQASVQVIDSTRCNADDAVQ 363
QY 362 GEVTEKMGAGTPEGVDTCQGDSSGGLMYQSDQMAHVIGVSKGTCGSPSTPGVYTKVS 421
DB 364 GEVTEKMGAGTPEGVDTCQGDSSGGLMYQSDQMAHVIGVSKGTCGSPSTPGVYTKVS 423
QY 422 AYLNMIYNNWKAEL 435
DB 424 AYLNMIYNNWKAEL 437

RESULT 12
ABU04936
ID ABU04936 standard; protein; 437 AA.
XX
AC ABU04936;
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1602.
XX
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KW protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.
XX

PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-APR-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicz RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.

Example 2; SEQ ID NO 1602; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SO Sequence 437 AA:

Query Match 99.8%; Score 2337; DB 6; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.6e-201;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSQPLNSLDVYKPLKRPRIPIIIMLSTASIIIVVILKVIDKYFL 61
DB 4 DPDSQPLNSLDVYKPLKRPRIPIIIMLSTASIIIVVILKVIDKYFL 63
QY 62 CGQPLHPIPRKOLCDGLDPLGDEBEHCYKSPPEGPAVAVRUSKDRSTLQVLDATGMW 121
DB 64 CGQPLHPIPRKOLCDGLDPLGDEBEHCYKSPPEGPAVAVRUSKDRSTLQVLDATGMW 123
QY 122 FSACFDNFTALAEATACROWGYSKPTFRAVEIGPDDLDVVEITENSQELRNRSSGRC 181
DB 124 FSACFDNFTALAEATACROWGYSKPTFRAVEIGPDDLDVVEITENSQELRNRSSGRC 183
QY 182 LSGSLVSLHCLACGSKSLKTPRVVVGGEASVDSWPMQVSIQYDKQHCYSGSILDPHVVLT 241
DB 184 LSGSLVSLHCLACGSKSLKTPRVVVGGEASVDSWPMQVSIQYDKQHCYSGSILDPHVVLT 243
QY 242 AHCFRKHITDVFNMKVRAAGSDKLSFSLAVAKIIIEFNPMYRKONDIALMKLQEPFLTS 301
DB 244 AHCFRKHITDVFNMKVRAAGSDKLSFSLAVAKIIIEFNPMYRKONDIALMKLQEPFLTS 303
QY 302 GTVRPICLPFDEBELTPATPLMTIIGWFTKONGKNSDILLQASVQVIDSTRCNADDAVQ 361
DB 304 GTVRPICLPFDEBELTPATPLMTIIGWFTKONGKNSDILLQASVQVIDSTRCNADDAVQ 363

```
ABU04920
ID ABU04920 standard; protein; 437 AA.
AC ABU04920;
XX
XX 29-JAN-2003 (first entry)
DE Human expressed protein tag (EPT) #1586.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KM receptor; transcription factor; cancer; MHC;
KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
FN WO200278524-A2.
PD 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOs INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1586; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 437 AA;
XX
XX Query Match 99.8%; Score 2337; DB 6; Length 437;
XX Best Local Similarity 100.0%; Pred. No. 1.6e-201;
XX Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 2 DPDSQPIINSLDVKRLRKPRIMETPRKVGPIIITALLSLASIIIVVLIKYLDKYFL 61
XX 4 DPDSQPIINSLDVKRLRKPRIMETPRKVGPIIITALLSLASIIIVVLIKYLDKYFL 63
XX
XX 62 CGOPLHFI PRKQLCDGELDCPLGDEDEHCVKSPFPGGPAVAVRLSKDRTLQVLD SATGNW 121
XX
XX
```

```
DB 64 CGOPLHFI PRKQLCDGELDCPLGDEDEHCVKSPFPGGPAVAVRLSKDRTLQVLD SATGNW 123
QY 122 FSACFDNFTEALAEACRQNGYSKPTFAVEIGPDOLDVVEITENSQELRMNSSGPC 181
DB 124 FSACFDNFTEALAEACRQNGYSKPTFAVEIGPDOLDVVEITENSQELRMNSSGPC 183
QY 182 LSGSLVSLHCLACGSKLKTFRVVGGEASVDSWFWQVSIQYDKQHVCGSILDPHWLTA 241
DB 184 LSGSLVSLHCLACGSKLKTFRVVGGEASVDSWFWQVSIQYDKQHVCGSILDPHWLTA 243
QY 242 AHCFRKTDFVFNKVRGSGKLSFSPSLAVAKIIIEFNMPYKDNIDALMKLQPLTFS 301
DB 244 AHCFRKTDFVFNKVRGSGKLSFSPSLAVAKIIIEFNMPYKDNIDALMKLQPLTFS 303
QY 302 GTVAPICLPFDEBELTPATPLMIIGWFTKQNGSKMSDILQASVQYIDSTRCADDAAYQ 361
DB 304 GTVAPICLPFDEBELTPATPLMIIGWFTKQNGSKMSDILQASVQYIDSTRCADDAAYQ 363
QY 362 GEVTERKMCAGIPFGVDTQGDGSGPLMYQSDGMHVGVISWVGCGGPGSTPGVYTKVS 421
DB 364 GEVTERKMCAGIPFGVDTQGDGSGPLMYQSDGMHVGVISWVGCGGPGSTPGVYTKVS 423
QY 422 AYLMWITVWKAEL 435
DB 424 AYLMWITVWKAEL 437
XX
XX RESULT 11
XX ABU04935
XX ID ABU04935 standard; protein; 437 AA.
XX
XX AC ABU04935;
XX
XX 29-JAN-2003 (first entry)
XX
XX Human expressed protein tag (EPT) #1601.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX Homo sapiens.
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOs INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1601; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor.
```

XX Sequence 435 AA;
 Query Match 99.8%; Score 2338; DB 6; Length 435;
 Best Local Similarity 99.8%; Pred. No. 1.3e-201;
 Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPSDQPLNSLDVYPLKRPRIEMETFRKVGIPRIIALLSLASIIIVVLIKVILDKYFL 60
 DB 1 MDPSDQPLNSLDVYPLKRPRIEMETFRKVGIPRIIALLSLASIIIVVLIKVILDKYFL 60
 QY 61 LCQGPLHPIPRKQJCDGLDCEPLGDEDEHCVKSPFEGPAAVAVRLSKRSTLQVLDATGN 120
 DB 61 LCQGPLHPIPRKQJCDGLDCEPLGDEDEHCVKSPFEGPAAVAVRLSKRSTLQVLDATGN 120
 QY 121 WFSACFDNFTALAEACRQWYSSKPTFRAVEIGPODDLDVVEITENSOELMRNNSGPG 180
 DB 121 WFSACFDNFTALAEACRQWYSSKPTFRAVEIGPODDLDVVEITENSOELMRNNSGPG 180
 QY 181 CLSGSLVSLHCLAGSKLTPRVVGGEEASVDSWPMQVSIQYDKOHVCGGSILDPHWLVT 240
 DB 181 CLSGSLVSLHCLAGSKLTPRVVGGEEASVDSWPMQVSIQYDKOHVCGGSILDPHWLVT 240
 QY 241 AHCGRKHTDVFNKAVAGSDKLSFPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 300
 DB 241 AHCGRKHTDVFNKAVAGSDKLSFPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 300
 QY 301 SGVVRPILCPFPDEBELTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADAY 360
 DB 301 SGVVRPILCPFPDEBELTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADAY 360
 QY 361 QGEVTEKMMGAGIPEGGVDTCCGDSGGLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKV 420
 DB 361 QGEVTEKMMGAGIPEGGVDTCCGDSGGLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKV 420
 QY 421 SAYLNMWYWKAEI 435
 DB 421 SAYLNMWYWKAEI 435

RESULT 9
 AAG79359 standard; protein; 437 AA.
 ID AAG79359
 AC AAG79359;
 XX
 DT 21-AUG-2002 (first entry)
 XX
 DE CUA8 preferred sequence.
 XX
 KW Colorectal cancer; CGA7; CUA8; modulating protein; screening;
 XX drug candidate; vaccine.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 36..63
 FT /note= "Transmembrane domain"
 PN US2002042067-A1.
 PD 11-APR-2002.
 PP 08-MAY-2001; 2001US-00851588.
 PR 17-AUG-2000; 2000US-00642252.
 PR 06-SEP-2000; 2000US-00656002.
 XX (MACK/) MACK D.
 PA (GISH/) GISH K C.
 PA (WILS/) WILSON K E.
 XX Mack D, Gish KC, Wilson KE;
 PI

XX WPI: 2002-453647/48.
 DR N-PSDB; AAT72976.
 XX
 PT Screening drug candidates for treating colorectal cancer, comprises
 PT determining the effect of the candidate on the expression profile gene of
 PT CGA7 or CUA8.
 XX
 PS Disclosure; Fig 8; 40pp; English.
 XX
 CC The sequences given in AAG79356-59 show the colorectal cancer proteins,
 CC CGA7 and CUA8. The CGA7 and CUA8 proteins are colorectal cancer
 CC modulating proteins and have been mapped to chromosomes 2 (CGA7) and 11
 CC (CUA8). These sequences may be used in the method of the invention for
 CC screening drug candidates. The method comprises adding a drug candidate
 CC to a cell that expresses an expression profile gene encoding CGA7, CUA8
 CC or fragments and determining the effect of the drug candidate on the
 CC expression of the expression profile gene. The new methods are used to
 CC screen bioactive agents for the ability to bind to or modulate the
 CC activity of CGA7 or CUA8 and evaluate the effect of a candidate
 CC colorectal cancer drug. An antibody to CGA7 or CUA8 can inhibit the
 CC activity of CGA7 or CUA8, respectively, and is used to screen for an
 CC agent that can interfere with the binding of CGA7 or CUA8 to the
 CC antibody. The antibody can be used to treat colorectal cancer. The
 CC colorectal cancer tissue, where the therapeutic group is a cytotoxic
 CC agent or a radioisotope. Antisense molecules are used to inhibit
 CC colorectal cancer in a cell. Nucleic acid segments encoding CGA7 or CUA8
 CC are used in a biochip. CGA7, CUA8 or a nucleic acid encoding it are used
 CC to elicit an immune response. CGA7 or CUA8 is used to determine the
 CC prognosis of an individual with colorectal cancer. Nucleic acid encoding
 CC CGA7 or CUA8 can be used in vaccines
 CC
 SQ Sequence 437 AA;
 Query Match 99.8%; Score 2337; DB 5; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.6e-201;
 Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DPDSDDQPLNSLDVYPLKRPRIEMETFRKVGIPRIIALLSLASIIIVVLIKVILDKYFL 61
 DB 4 DPDSDDQPLNSLDVYPLKRPRIEMETFRKVGIPRIIALLSLASIIIVVLIKVILDKYFL 63
 QY 62 CGGPLHPIPRKQJCDGLDCEPLGDEDEHCVKSPFEGPAAVAVRLSKRSTLQVLDATGN 121
 DB 64 CGGPLHPIPRKQJCDGLDCEPLGDEDEHCVKSPFEGPAAVAVRLSKRSTLQVLDATGN 123
 QY 122 FSACFDNFTALAEACRQWYSSKPTFRAVEIGPODDLDVVEITENSOELMRNNSGPG 181
 DB 124 FSACFDNFTALAEACRQWYSSKPTFRAVEIGPODDLDVVEITENSOELMRNNSGPG 183
 QY 182 LSGSLVSLHCLAGSKLTPRVVGGEEASVDSWPMQVSIQYDKOHVCGGSILDPHWLVT 241
 DB 184 LSGSLVSLHCLAGSKLTPRVVGGEEASVDSWPMQVSIQYDKOHVCGGSILDPHWLVT 243
 QY 242 AHCGRKHTDVFNKAVAGSDKLSFPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 301
 DB 244 AHCGRKHTDVFNKAVAGSDKLSFPSLAVAKIIIEFNPMYPKNDIALMLKQPLTF 303
 QY 302 GTVVRPILCPFPDEBELTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADAY 361
 DB 304 GTVVRPILCPFPDEBELTPATPLMIIGWFTKONGKMSDILLQASVQVIDSTRCNADAY 363
 QY 362 GEVTEKMMGAGIPEGGVDTCCGDSGGLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKV 421
 DB 364 GEVTEKMMGAGIPEGGVDTCCGDSGGLMYQSDQHWVGIYSWGYCGGSPSTPGVYTKV 423
 QY 422 AYLNWYWKAEI 435
 DB 424 AYLNWYWKAEI 437

RESULT 10

```

FT Modified-site /note= "protein kinase C phosphorylation site"
FT 303
FT Modified-site /note= "protein kinase C phosphorylation site"
FT 351
FT Modified-site /note= "protein kinase C phosphorylation site"
FT 360
FT Modified-site /note= "tyrosine kinase phosphorylation site"
FT 365
FT Active-site /note= "protein kinase C phosphorylation site"
FT 385
FT /note= "characteristic of serine protease"
XX
XX W09936550-A2.
XX
XX 22-JUL-1999.
XX
XX 12-JAN-1999; 99WO-US000655.
XX
XX 16-JAN-1998; 98US-00008271.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bardman O, Hillman JL, Yue H, Guegler KJ, Corley NC, Tang YT,
XX Shah P;
XX WPI; 1999-430616/36.
XX DR N-PSDB; AAX87154.
XX
XX Novel human protease molecules useful in the treatment of developmental
XX disorders and/or cancers.
XX
XX Claim 1; Page 74-75; 90pp; English.
XX
XX The present sequence represents novel human protease HUPM-6, as deduced
XX from the consensus sequence (see AAX87154) of overlapping cDNA clones
XX obtained from various libraries. Northern analysis shows expression of
XX HUPM-6 in gastrointestinal, and male and female reproductive cDNA
XX libraries. Approximately 65% of these libraries are associated with
XX neoplastic disorders, and 22% with the immune response. The invention
XX provides 12 new human proteases, 1.e. HUPM-1 to -12 (see AAY06432-43),
XX and the polynucleotides encoding them (see AAX87149-60). Also provided
XX are vectors, host cells and methods for producing HUPM polypeptides, as
XX well as agonists and antagonists of HUPM. Methods for treating or
XX preventing cell proliferative disorders and immune disorders using HUPM
XX or HUPM antagonists are claimed
XX
XX Sequence 435 AA:
SQ
Query Match 99.8%; Score 2338; DB 2; Length 435;
Best Local Similarity 99.8%; Pred. No. 1.3e-201;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVKKPLRKPRIPMETPRKVGIPITIIALSLIIIVVLKVLDDKYF 60
DB 1 MDPDSQPLNSLDVKKPLRKPRIPMETPRKVGIPITIIALSLIIIVVLKVLDDKYF 60
QY 61 LCGGPLHFIPIKQICDGLDCLGDEDEBHCYKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
DB 61 LCGGPLHFIPIKQICDGLDCLGDEDEBHCYKSPPEGPAVAVRLSKORSTLQVLDSATGN 120
QY 121 WFSACFNFTFALAEATACRQMGYSKPTFAVEIGEPDODLDVEITENSQELRNRNSGP 180
DB 121 WFSACFNFTFALAEATACRQMGYSKPTFAVEIGEPDODLDVEITENSQELRNRNSGP 180
QY 181 CLASGLVSLHCLACGSKSLKTPRVVGGEEASVDSWPMOVSIOYDRQHVCGGSIIDPHWVLT 240
DB 181 CLASGLVSLHCLACGSKSLKTPRVVGGEEASVDSWPMOVSIOYDRQHVCGGSIIDPHWVLT 240
QY 241 AAHCFRKHITDVFNKVRAGSDKLGSPSLAAVAKIIIIIEFNMYFRKNDIAMLKQLPPLTF 300
DB 241 AAHCFRKHITDVFNKVRAGSDKLGSPSLAAVAKIIIIIEFNMYFRKNDIAMLKQLPPLTF 300
QY 301 SGTVRPICLPFDEBLPATPLMIIGMGFTKONGGKMSDIILOASVOYIDSTRCNADDAY 360

```

```

DB 301 SGTVRPICLPFDEBLPATPLMIIGMGFTKONGGKMSDIILOASVOYIDSTRCNADDAY 360
QY 361 QGEVTERKMCAGIPREGVDTCQSGGGLPMYQSDQMHVGVISNGYGGCGSGSTGVYTKV 420
DB 361 QGEVTERKMCAGIPREGVDTCQSGGGLPMYQSDQMHVGVISNGYGGCGSGSTGVYTKV 420
QY 421 SAYLMIYVWKAEL 435
DB 421 SAYLMIYVWKAEL 435
RESULT 8
ABU04930
ID ABU04930 standard; protein; 435 AA.
AC ABU04930;
XX
XX 29-JAN-2003 (first entry)
XX
XX Human expressed protein tag (EPT) #1596.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.
XX
XX Homo sapiens.
XX
XX W0200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
XX 21-MAY-2001; 2001US-0292544P.
XX 08-AUG-2001; 2001US-0310801P.
XX 01-OCT-2001; 2001US-0326370P.
XX 04-DEC-2001; 2001US-0336780P.
XX 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOX INC.
XX
XX Chicx RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1596; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX lymphoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX myeloma or leukemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

```


PR 01-OCT-2001; 2001US-0326370P.
 PR 04-DEC-2001; 2001US-0336780P.
 PR 20-FEB-2002; 2002US-0358985P.
 XX
 XX (ZYCO-) ZYCOs INC.
 PA
 PI Chicx RM, Tomlinson AJ, Urban RG;
 XX WPI; 2003-040607/03.
 XX
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
 PT cytoskeletal proteins, receptors or transcription factors), useful for
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
 PT leukemia.
 XX
 XX
 PS Example 2; SEQ ID NO 1578; 134pp; English.
 XX

CC The invention describes a purified polypeptide, which comprises a
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,
 CC transporter, cytoskeletal protein, receptor or transcription factor. The
 CC polypeptide is useful as an immunogenic composition for eliciting in a
 CC mammal an immunogenic response directed against any of the purified
 CC polypeptide. The purified polypeptide, or the antibody that binds to this
 CC polypeptide, is useful for treating cancer. The polypeptide is also
 CC useful for identifying compounds that binds to a naturally processed
 CC class I or class II MHC-binding polypeptide. The polypeptides and
 CC polynucleotides are particularly useful for treating or preventing
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
 CC lymphoma or leukemia. These are also useful for screening agents for
 CC treating the above mentioned diseases. This sequence represents an
 CC expressed protein tag (EPT) isolated from human tissue for translational
 CC profiling. Note: This sequence does not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX

SO Sequence 435 AA;

Query Match 100.0%; Score 2342; DB 6; Length 435;
 Best Local Similarity 100.0%; Pred. No. 5.8e-202;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVYKPRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVIIDKYYF 60
 DB 1 MDPDSQPLNSLDVYKPRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVIIDKYYF 60
 QY 61 LCGQPLHPIPRKQICDELCPLEGDEBEHCVSFPEEPAVAVRISKDSTLQVDSATGN 120
 DB 61 LCGQPLHPIPRKQICDELCPLEGDEBEHCVSFPEEPAVAVRISKDSTLQVDSATGN 120
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDXQHVCGGSIIDPHWLT 240
 DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDXQHVCGGSIIDPHWLT 240
 QY 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 QY 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 QY 301 SGTVRPILCPFDEDELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360
 DB 301 SGTVRPILCPFDEDELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360
 QY 361 QGEVTEKMGACIGPEGVDTCQSDSGPPLMYQSDQMHVVGIVSWGCGGSGSTPEVYTKV 420
 DB 361 QGEVTEKMGACIGPEGVDTCQSDSGPPLMYQSDQMHVVGIVSWGCGGSGSTPEVYTKV 420
 QY 421 SAYLNMIVNWKAEL 435
 DB 421 SAYLNMIVNWKAEL 435
 QY 421 SAYLNMIVNWKAEL 435
 DB 421 SAYLNMIVNWKAEL 435

RESULT 5
 ADU91838
 ID ADU91838 standard; protein; 461 AA.
 XX
 XX AC ADU91838;
 XX
 XX DT 24-FEB-2005 (first entry)
 XX
 XX DE Ovr115.
 XX
 XX KM cytosstatic; Ovr115.
 XX
 XX OS Homo sapiens.
 OS Tobacco etch virus.
 OS Unidentified.
 XX
 XX PN WO2004104173-A2.
 XX
 XX PD .02-DEC-2004.
 XX
 XX PF 17-MAY-2004; 2004WO-US015258.
 XX
 XX PR 16-MAY-2003; 2003US-0471068P.
 PR 05-APR-2004; 2004US-0559730P.
 XX
 XX PA (DIAD-) DIADEXUS INC.
 XX
 XX PI Pilkington G, Keller G, Li W, Corral L, Simon I;
 XX WPI; 2004-834291/82.
 XX

PT New isolated anti-ovarian, pancreatic, lung or breast cancer antigen
 PT (Ovr115) antibody that binds to Ovr115 on a mammalian cell, useful for
 PT diagnosing or treating ovarian, pancreatic and colon cancer.
 PT
 XX
 XX

Example 1; SEQ ID NO 3; 169pp; English.

CC The invention relates to an isolated anti-ovarian, pancreatic, lung or
 CC breast cancer antigen (Ovr115) antibody that binds to Ovr115 on a
 CC mammalian cell in vivo. The methods and compositions of the present
 CC invention are useful for producing anti-Ovr115 antibody compositions for
 CC diagnosing or treating Ovr115-expressing ovarian, pancreatic and colon
 CC cancer cells. The present sequence represents the amino acid sequence of
 CC Ovr115.
 XX
 XX

SO Sequence 461 AA;

Query Match 100.0%; Score 2342; DB 8; Length 461;
 Best Local Similarity 100.0%; Pred. No. 6.2e-202;
 Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVYKPRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVIIDKYYF 60
 DB 1 MDPDSQPLNSLDVYKPRKPRIPMETFRKVGIPITIALSLASIIIVVLIKVIIDKYYF 60
 QY 61 LCGQPLHPIPRKQICDELCPLEGDEBEHCVSFPEEPAVAVRISKDSTLQVDSATGN 120
 DB 61 LCGQPLHPIPRKQICDELCPLEGDEBEHCVSFPEEPAVAVRISKDSTLQVDSATGN 120
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 DB 121 WFSACFDFNFTALAEACRQWYSSKPTFRAVEIGPPDODLDVVEITENSQELRMNNSGP 180
 QY 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDXQHVCGGSIIDPHWLT 240
 DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGGEEASVDSWPQVSIQDXQHVCGGSIIDPHWLT 240
 QY 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 QY 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 DB 241 AAHCFRKTIDVFNKVRAGSDKLSFPSLAIVAKIIIFENPMYPRKNDIALMKLOPPLTF 300
 QY 301 SGTVRPILCPFDEDELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360
 DB 301 SGTVRPILCPFDEDELTPATPLMIIGWGFTKONGKMSDILLQASVOYIDSTRCNADAY 360

```

OY 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
DB 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
OY 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
DB 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
OY 361 QGEVTERKMKACGIPGSGVDTCQDSDSGPLMYQSDQMHVGIYSWGIGCGSPTPGVYTKV 420
DB 361 QGEVTERKMKACGIPGSGVDTCQDSDSGPLMYQSDQMHVGIYSWGIGCGSPTPGVYTKV 420
OY 421 SAYLNMWYVWKAEL 435
DB 421 SAYLNMWYVWKAEL 435

RESULT 3
ABU04931
ID ABU04931 standard; protein; 435 AA.
XX
XX ABU04931;
AC
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1597.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX PA
XX PI Chicx RM, Tomlinson AJ, Urban RG;
XX DR WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukemia.
XX
XX Example 2; SEQ ID NO 1597; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide is also
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

```

```

CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 435 AA;
Query Match 100.0%; Score 2342; DB 6; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MDPDSQPLNSLDVPLRKPRIPMETPRKVGIPPIIALSLASIIIVVLIKVIDKYFF 60
DB 1 MDPDSQPLNSLDVPLRKPRIPMETPRKVGIPPIIALSLASIIIVVLIKVIDKYFF 60
OY 61 ICGQPLHFIPIRQQLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDNSTQLVDSATGN 120
DB 61 ICGQPLHFIPIRQQLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDNSTQLVDSATGN 120
OY 121 WFSACFDNFTETALATACROWGYSKPTFRAVEIGPDQDLVETITENSQELRMNSSGP 180
DB 121 WFSACFDNFTETALATACROWGYSKPTFRAVEIGPDQDLVETITENSQELRMNSSGP 180
OY 181 CLSGSLVSLHCLACGSKSLKTPRVVGEBASVDSWPMQVSIQYDKOHVCGSILDPHWVLT 240
DB 181 CLSGSLVSLHCLACGSKSLKTPRVVGEBASVDSWPMQVSIQYDKOHVCGSILDPHWVLT 240
OY 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
DB 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
OY 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
DB 241 AAHCRKHTDVFNWKVRAGSDKLGSPSLAVAKIIIEFNPMYPKNDIALMKLOPLTF 300
OY 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
DB 301 SGTVRPICPFPEDELTPTPLMIIGWFTKONGKMSDILLQASVQVIDSTRCANADAY 360
OY 361 QGEVTERKMKACGIPGSGVDTCQDSDSGPLMYQSDQMHVGIYSWGIGCGSPTPGVYTKV 420
DB 361 QGEVTERKMKACGIPGSGVDTCQDSDSGPLMYQSDQMHVGIYSWGIGCGSPTPGVYTKV 420
OY 421 SAYLNMWYVWKAEL 435
DB 421 SAYLNMWYVWKAEL 435

RESULT 4
ABU04912
ID ABU04912 standard; protein; 435 AA.
XX
XX ABU04912;
AC
XX
DT 29-JAN-2003 (first entry)
XX
DE Human expressed protein tag (EPT) #1578.
XX
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
KM protease; protease inhibitor; transporter; cytoskeletal protein;
KW receptor; transcription factor; cancer; MHC;
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX Homo sapiens.
XX
XX WO200278524-A2.
XX
XX 10-OCT-2002.
XX
XX 28-MAR-2002; 2002WO-US009671.
XX
XX 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.

```


CC and diagnosing diseases such as arthritis, chronic obstructive pulmonary
CC disorder (COPD), cancer, osteoporosis, aberrant wound healing,
CC angiogenesis, inflammatory disorders, diabetes, stroke and cardiovascular
CC diseases. Seripancrin gene are useful in chromosome localization
CC studies, as tools for tissue expression studies and also in gene therapy.
CC The polypeptides of the invention are used for identifying agonists and
CC antagonists useful for treating conditions associated with seripancrin
CC imbalance. These polypeptides are also useful as vaccines. The present
CC sequence is seripancrin protein. The seripancrin gene is located on human
CC chromosome 11q22-q23
XX
XX Sequence 435 AA;
SQ
Query Match 100.0%; Score 2342; DB 4; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
Db 1 MDPDSQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
QY 61 LCGPPLHPIPRKQUCDGLDPLGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
Db 61 LCGPPLHPIPRKQUCDGLDPLGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
QY 121 WFSACFDNFTALAEACRQWGSKPTFRFAVEIGPDODLVVEITENSQELRMNSGP 180
Db 121 WFSACFDNFTALAEACRQWGSKPTFRFAVEIGPDODLVVEITENSQELRMNSGP 180
QY 181 CLSGSLVSLHCLACGSKLKTFRVVGEEASVDSMPWQVSIQYDKQHVCGGSLDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGEEASVDSMPWQVSIQYDKQHVCGGSLDPHWLT 240
QY 241 AAHCFRKATDVFNMKVRAGSDKLSFSLAVAKIIIEFNPMYPRKNDIALMKLQFPLTF 300
Db 241 AAHCFRKATDVFNMKVRAGSDKLSFSLAVAKIIIEFNPMYPRKNDIALMKLQFPLTF 300
QY 301 SGTARPLLPFDELTATPLMTIIGKGFTRQNGKMDIILQASVOYIDSTRCADAY 360
Db 301 SGTARPLLPFDELTATPLMTIIGKGFTRQNGKMDIILQASVOYIDSTRCADAY 360
QY 361 QGEVTEKMKACAGIPREGVDTCCGDSGGLPMYQSDMHVVGIVSMWGGCGSTPGVYTKV 420
Db 361 QGEVTEKMKACAGIPREGVDTCCGDSGGLPMYQSDMHVVGIVSMWGGCGSTPGVYTKV 420
QY 421 SAYLNIWYVWKAEL 435
Db 421 SAYLNIWYVWKAEL 435
RESULT 2
AAG78577 ID AAG78577 standard; protein; 435 AA.
XX
XX AAG78577;
AC
DT 07-MAY-2002 (first entry)
XX
XX Human serine protease D-G amino acid sequence.
DE
XX Serine protease; D-G; human; zymogen; enzyme; cytosolic;
XX antiinflammatory; dermatological; anticoagulation; cancer; skin disorder;
KW neuropathic pain; inflammatory disorder; coagulation diathesis;
KW thrombosis; laundry detergent; skin care; gene therapy.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Region 31..52
FT Active-site /note="hydrophobic transmembrane domain"
FT 202..203
FT Active-site /note="residues spanning the zymogen cleavage site"
FT 243

FT /note= "catalytic triad residue"
FT Active-site 339
FT /note= "catalytic triad residue"
FT Msc-difference 361
FT /note= "Encoded by CTG"
FT Active-site 385
FT /note= "catalytic triad residue"
XX
XX WO200202011-A1.
XX
XX 10-JAN-2002.
XX
XX 08-JUN-2001; 2001WO-US018568.
XX
XX 30-JUN-2000; 2000US-00607745.
XX
XX (ORTH) ORTHO-MCNEIL PHARM INC.
XX Darrow AL, Qi J, Andrade-Gordon P;
XX WPI; 2002-106601/14.
XX N-PSDB; AAI64284.
XX
XX Nucleic acid encoding a serine protease called D-G protein which is
XX useful for identifying modulators that are useful for treating a
XX condition which is mediated by protease D-G, e.g. cancer, skin disorders,
XX or neuropathic pain.
XX
XX Claim 13; Fig 1B; 81pp; English.
XX
XX The invention relates to an isolated and purified nucleic acid that
XX encodes a serine protease called D-G protein. The activity of the protein
XX of the invention may be described as cytosolic, antiinflammatory,
XX dermatological and anticoagulation. The serine protease of the invention
XX is a member of the trypsin/chymotrypsin-like (S1) serine protease family,
XX which play an important role in processes such as digestion and
XX regulatory amplification cascades through the proteolytic activation of
XX inactive zymogen precursors. Protease D-G modulating compounds are useful
XX for treating a condition which is mediated by protease D-G, e.g. cancer,
XX skin disorders, neuropathic pain, inflammatory disorders, or coagulation
XX diathesis/thrombosis. The polynucleotide encoding the protease is useful
XX for identifying, detecting or isolating mutant forms of DNA molecules
XX encoding the protease. The protease is useful for identifying modulators
XX of the functional protease. The D-G protein can be used for formulation
XX of compositions for laundry detergents and skin care products. Protease D
XX -G gene therapy may be used to introduce protease D-G into the cells of
XX target organisms. As the D-G protein is derived from a human, it is less
XX likely to produce an allergic reaction in sensitive individuals when used
XX in formulations for laundry detergents and skin care products. The
XX current sequence represents the human serine protease D-G amino acid
XX sequence
XX
XX Sequence 435 AA;
SQ
Query Match 100.0%; Score 2342; DB 5; Length 435;
Best Local Similarity 100.0%; Pred. No. 5.8e-202;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDPDSQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
Db 1 MDPDSQPLNSLDVYKPKRPRIPIIALLSLASIIIVVLIKVLIDKYYF 60
QY 61 LCGPPLHPIPRKQUCDGLDPLGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
Db 61 LCGPPLHPIPRKQUCDGLDPLGEDEBHCYKSPFEGPAVAVRLSKDRSTIQVDSATGN 120
QY 121 WFSACFDNFTALAEACRQWGSKPTFRFAVEIGPDODLVVEITENSQELRMNSGP 180
Db 121 WFSACFDNFTALAEACRQWGSKPTFRFAVEIGPDODLVVEITENSQELRMNSGP 180
QY 181 CLSGSLVSLHCLACGSKLKTFRVVGEEASVDSMPWQVSIQYDKQHVCGGSLDPHWLT 240
Db 181 CLSGSLVSLHCLACGSKLKTFRVVGEEASVDSMPWQVSIQYDKQHVCGGSLDPHWLT 240

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:41:18 ; Search time 188 Seconds

(without alignments)
1016.648 Million cell updates/sec

Title: US-10-803-530-2

Sequence: 1 MDPDSQDPLNSLDVVKRLAKP.....VYTVSAVLMYVWKAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseq_21:*
2: Geneseqp1980s:*
3: Geneseqp1990s:*
4: Geneseqp2000s:*
5: Geneseqp2001s:*
6: Geneseqp2002s:*
7: Geneseqp2003as:*
8: Geneseqp2003bs:*
9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	435	4 AAY72558	Aay72558 Human ser
2	2342	100.0	435	5 AAG78577	Aag78577 Human ser
3	2342	100.0	435	6 ABU04931	Abu04931 Human exp
4	2342	100.0	435	6 ABU04912	Abu04912 Human exp
5	2342	100.0	461	8 ADU91838	Adu91838 Ovr115. 2
6	2342	100.0	461	9 AEA10749	Aea10749 Mutated h
7	2338	99.8	435	2 AAY06437	Aay06437 Human pro
8	2338	99.8	435	6 ABU04930	Abu04930 Human exp
9	2337	99.8	437	5 AAG79359	Aag79359 CVA8 prei
10	2337	99.8	437	6 ABU04920	Abu04920 Human exp
11	2337	99.8	437	6 ABU04935	Abu04935 Human exp
12	2337	99.8	437	6 ABU04936	Abu04936 Human exp
13	2337	99.8	437	6 ABU04915	Abu04915 Human exp
14	2337	99.8	437	7 ADBR0525	Adbr0525 Human exp
15	2337	99.8	437	7 ADN39878	Adn39878 Ovarian c
16	2337	99.8	437	7 ADN39513	Adn39513 Cancer/an
17	2337	99.8	437	7 ADN39461	Adn39461 Cancer/an
18	2337	99.8	437	7 ADN39473	Adn39473 Cancer/an
19	2337	99.8	437	8 ADN04864	Adn04864 Antipsori
20	2337	99.8	437	8 ADU25740	Adu25740 Human TMP
21	2337	99.8	437	9 ADW04409	Adw04409 Human Ova
22	2337	99.8	437	9 AEA54849	Aea54849 Human tum
23	2333	99.6	435	8 ABO84561	Ab084561 Human can
24	2329	99.4	492	4 AAY72559	Aay72559 Human ser

25	2329	99.4	492	6 ABU04913	Abu04913 Human exp
26	2329	99.4	492	6 ABU04932	Abu04932 Human exp
27	2328	99.4	437	8 ABO84562	Ab084562 Human can
28	2319	99.0	437	4 AAE06931	Aae06931 Human mem
29	2319	99.0	437	6 ABU04917	Abu04917 Human exp
30	2319	99.0	437	6 ABU04922	Abu04922 Human exp
31	2319	99.0	437	7 AD110373	Ad110373 Human cel
32	2319	99.0	437	7 AD1346897	Ad1346897 Human cel
33	2297.5	98.1	432	3 AAY99417	Aay99417 Human tra
34	2297.5	98.1	432	4 AAB66166	Aab66166 Protein o
35	2297.5	98.1	432	4 AAB87581	Aab87581 Human PRO
36	2297.5	98.1	432	5 AAB87581	Aab87581 Human PRO
37	2297.5	98.1	432	5 AAB87581	Aab87581 Human PRO
38	2297.5	98.1	432	5 AAB87581	Aab87581 Human PRO
39	2297.5	98.1	432	5 AAB87581	Aab87581 Human PRO
40	2297.5	98.1	432	6 ABU58564	Abu58564 Human sec
41	2297.5	98.1	432	6 ABU84427	Abu84427 Human PRO
42	2297.5	98.1	432	6 ABU84427	Abu84427 Human sec
43	2297.5	98.1	432	6 ABR65301	Ab65301 Human sec
44	2297.5	98.1	432	6 ABR65301	Ab65301 Human sec
45	2297.5	96.1	432	6 ABU82870	Abu82870 Human PRO

ALIGNMENTS

RESULT 1
AAY72558
ID AAY72558 standard; protein, 435 AA.

AC AAY72558;
XX

DT 02-MAY-2001 (first entry)
XX

DE Human seripancrin protein.
XX

XX Human; seripancrin; serine protease; chromosome 11q22-q23; therapy;
XX arthritits; chronic obstructive pulmonary disorder; COPD; cancer;
KW osteoporosis; aberrant wound healing; angiogenesis; diabetes;
KW inflammatory disorder; stroke; cardiovascular disease; gene therapy;
KW vaccine; cytostatic; cerebroprotective; vulnerary; osteopathic.

XX Homo sapiens.
OS

XX WO200104141-A2.
PN

XX 18-JAN-2001.
PD

XX 04-JUL-2000; 2000WO-EP006211.
PF

XX 12-JUL-1999; 99EP-00113428.
PR

XX (MERCK) MERCK PATENT GMBH.
PA

XX Suendermann B, Hofmann U, Matzku S, Wilbert O;
PI

XX WPI; 2001-147177/15.
DR

XX N-PSDB; AAD02556.
DR

XX New extracellular serine protease Seripancrin, useful for treating
PT cancer, osteoporosis, arthritits, chronic obstructive pulmonary disease,
PT diabetes, inflammatory disorders, stroke, angiogenesis and aberrant wound
PT healing.

XX Claim 2; Page 39-40; 45pp; English.
PS

XX The present invention relates to seripancrin polynucleotides, and
CC polypeptides encoded by them. Seripancrin are members of serine protease
CC family. This protein contains a transmembrane domain, a low density
CC lipoprotein (LDL) domain, protease domain and a scavenger receptor
CC cysteine-rich (SRCR) domain. The LDL and SRCR domains help to define the
CC specificity of seripancrin's intra and intermolecular interactions. The
CC polynucleotides and polypeptides of the invention are useful for treating

```
RESULT 14
US-09-851-588-6
; Sequence 6, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; TITLE OF INVENTION: OF SCREENING FOR COLORECTAL CANCER MODULATORS
; FILE REFERENCE: A-68829-1/DB/JUD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-6

Query Match          93.4%; Score 2188; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 6.1e-215;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 VGPIPIIILSLASIIIVVLIKVLIDKYFLACGQPLFIPIRKQDCDELDCPLGEDEEH 89
DB 1 VGPIIILSLASIIIVVLIKVLIDKYFLACGQPLFIPIRKQDCDELDCPLGEDEEH 60

QY 90 CVKSPFEPGPAVAVRLSKDRSTLOVLDSATGNMFACPDNFTALAEIACROWGYSSKPTF 149
DB 61 CVKSPFEPGPAVAVRLSKDRSTLOVLDSATGNMFACPDNFTALAEIACROWGYSSKPTF 120

QY 150 RAVEIGPDQDLVVEITENSQELMRNNSGPGCLSGSLVSLHCLACGSKLTPRVVGGEBA 209
DB 121 RAVEIGPDQDLVVEITENSQELMRNNSGPGCLSGSLVSLHCLACGSKLTPRVVGGEBA 180

QY 210 SVDSFPMQVSIQYDKQHVCGGSIIDPHWVLTAAHCFRKHDTVENMKVRAAGSDKLSFPSL 269
DB 181 SVDSFPMQVSIQYDKQHVCGGSIIDPHWVLTAAHCFRKHDTVENMKVRAAGSDKLSFPSL 240

QY 270 AVAKIIIEFNMYPKNDIALMKLOPLTFSGTVRPICLPFDEELTPATPLMIIGMF 329
DB 241 AVAKIIIEFNMYPKNDIALMKLOPLTFSGTVRPICLPFDEELTPATPLMIIGMF 300

QY 330 TKONGGKMSDILLQASVOVLDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCGDSGGPL 389
DB 301 TKONGGKMSDILLQASVOVLDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCCGDSGGPL 360

QY 390 MYQSDQMHVVGIVSWGYGGGSPSTPGVYTKVSAYLNMWYVWKAEL 435
DB 361 MYQSDQMHVVGIVSWGYGGGSPSTPGVYTKVSAYLNMWYVWKAEL 406

RESULT 15
US-09-607-745-9
; Sequence 9, Application US/09607745
; Patent No. 6750034
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew L
; APPLICANT: Qi, Jain-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 9
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
US-09-607-745-9

Query Match          54.1%; Score 1266; DB 2; Length 292;
Best Local Similarity 97.4%; Pred. No. 7.1e-121;
Matches 228; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 202 RVVGGEEASVDSWPMQVSIQYDKQHVCGGSIIDPHWVLTAAHCFRKHDTVENMKVRAAGSD 261
DB 51 KIVGGVALDVDSWPMQVSIQYDKQHVCGGSIIDPHWVLTAAHCFRKHDTVENMKVRAAGSD 110

QY 262 KLSGFPSELAVAKIIIEFNMYPKNDIALMKLOPLTFSGTVRPICLPFDEELTPATP 321
DB 111 KLSGFPSELAVAKIIIEFNMYPKNDIALMKLOPLTFSGTVRPICLPFDEELTPATP 170

QY 322 LMIIGMGFTKONGGKMSDILLQASVOVLDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCC 381
DB 171 LMIIGMGFTKONGGKMSDILLQASVOVLDSTRCNADDAVQGEVTEKMCAGIPEGGVDTCC 230

QY 382 QGDSGGPLMYQSDQMHVVGIVSWGYGGGSPSTPGVYTKVSAYLNMWYVWKAEL 435
DB 231 QGDSGGPLMYQSDQMHVVGIVSWGYGGGSPSTPGVYTKVSAYLNMWYVWKAEL 284

Search completed: December 5, 2005, 13:54:42
Job time : 47 secs
```

Query Match	98.1%	Score 2297.5	DB 2	Length 432
Best Local Similarity	98.8%	Pred. No. 4,36-226		
Matches 429	Conservative	0	Mismatches 0	Indels 5
			Gaps	1
QY	2	DPDSQPINSLDVKPRKPRIPMETFRKRGPIIITALLSLASIIIVWLIVKILTDKKYFL	61	
Db	4	DPDSQPINSLDVKPRKPRIPMETFRKRGPIIITALLSLASIIIVWLIVKILTDKKYFL	63	
QY	62	CGQPLHPIPRKQLCDGELDCPLGDBDEHCVKSPFEGPAVAARLSKDRSTLIQVLSATGNW	121	
Db	64	CGQPLHPIPRKQLCDGELDCPLGDBDEHCVKSPFEGPAVAARLSKDRSTLIQVLSATGNW	123	
QY	122	FSACFDNFTMALAETACRQMGYSKPTFAVEIGPDQLDVVEITENSQELRMNRSGPC	181	
Db	124	FSACFDNFTMALAETACRQMGYS-----RAVEIGPDQLDVVEITENSQELRMNRSGPC	178	
QY	182	LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSMPQVSIQYDKQAVCGGSLIDPHWVTLTA	241	
Db	179	LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSMPQVSIQYDKQAVCGGSLIDPHWVTLTA	238	
QY	242	AHCRKRTDVENMKVRAGSDFLGSPPSLAAVKIIITFENPMYPRKNDIATLAKLQPLPLFS	301	
Db	239	AHCRKRTDVENMKVRAGSDFLGSPPSLAAVKIIITFENPMYPRKNDIATLAKLQPLPLFS	298	

Query Match	93.8%;	Score 2196;	DB 2;	Length 423;
Best Local Similarity	97.1%;	Pred. No. 9.9e-216;		
Matches 408;	Conservative 1;	Mismatches 11;	Indels 0;	Gaps 0;

QY	16	PLKRPRIEMETFRKVGIPITIIIALSLASIIIVVLLIKIILDKYFELCGQPLHPIPRKOLC	75
Db	4	PCANPVSFWMRSESVGPIPLIITALLSLASIIIVVLLIKIILDKYFELCGQPLHPIPRKOLC	63
QY	76	DGEILDCLGDEDEHCYVSFPBGPAVAVRLSKNSTLOYLSATGNWFSACDNFTALAE	135
Db	64	DGEILDCLGDEDEHCYVSFPBGPAVAVRLSKNSTLOYLSATGNWFSACDNFTALAE	123
QY	136	TACROMGYSSKPTFRAVEIGPDODLDAVEITENSGELRMRSSGPGCLSGSLVSLHCLACG	195
Db	124	TACROMGYSSKPTFRAVEIGPDODLDAVEITENSGELRMRSSGPGCLSGSLVSLHCLACG	183
QY	196	KSLLKTPRVVGGEEASVDSWPMQVSIQYDKOHVCGSILDPHMVLTAACHCFKHTDVFENWK	255
Db	184	KSLLKTPRVVGGEEASVDSWPMQVSIQYDKOHVCGSILDPHMVLTAACHCFKHTDVFENWK	243
QY	256	VRAGSDDLGSFPLSAVAKIIIEENPMYRPANDNIALMKLOEPLTFSGTVRRIICLPFDEE	315
Db	244	VRAGSDDLGSFPLSAVAKIIIEENPMYRPANDNIALMKLOEPLTFSGTVRRIICLPFDEE	303
QY	316	LTPATPLMIIGWGTGKNGGKMSDILLQASVOVLDSTRCANADAYQGEVTEKMKACAGIPE	375
Db	304	LTPATPLMIIGWGTGKNGGKMSDILLQASVOVLDSTRCANADAYQGEVTEKMKACAGIPE	363
QY	376	GGVDTQCGDSGGPLMTQSDMHWVVGIVSMGYGCGGSESTPGVYTTVTSAYLMTINVMKAEI	435
Db	364	GGVDTQCGDSGGPLMTQSDMHWVVGIVSMGYGCGGSESTPGVYTTVTSAYLMTINVMKAEI	423

Oy		302	GTVRICKLPFPEBELTPATPLMIIGMFTKONGKKMSDILLQASVOVIDSTCNADDAVQ	361
Dd		299	GTVRIICLPFDEBELTPATPLMIIGMFTKONGKKMSDILLQASVOVIDSTCNADDAVQ	358
Oy		362	GEVERKMVCAGIPGEGVDTCQSDSGPPLMYOSDOMHWIVGISWVGCGGSPITPGVTYTKVS	421
Dd		359	GEVERKMVCAGIPGEGVDTCQSDSGPPLMYOSDOMHWIVGISWVGCGGSPITPGVTYTKVS	418
Oy		422	AYLNMIVYWKAEL 435	
Dd		419	AYLNMIVYWKAEL 432	
RESULT 8				
US-10-015-671A-275				
; Sequence 275, Application US/10015671A				
; Patent No. 6946263				
; GENERAL INFORMATION:				
; APPLICANT: Baker, Kevin P.				
; APPLICANT: Botstein, David				
; APPLICANT: Desnoyers, Luc				
; APPLICANT: Baton, Dan I.				
; APPLICANT: Ferrara, Napoleone				
; APPLICANT: Fong, Sherman				
; APPLICANT: Gao, Wei-Qiang				
; APPLICANT: Goddard, Audrey				
; APPLICANT: Godowski, Paul J.				
; APPLICANT: Grimaldi, Christopher J.				
; APPLICANT: Gurney, Austin L.				
; APPLICANT: Hillan, Kenneth J.				
; APPLICANT: Pan, James				
; APPLICANT: Paoni, Nicholas F.				
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic				
; TITLE OF INVENTION: Acids Encoding the Same				
; FILE REFERENCE: P2830PIC47				
; CURRENT APPLICATION NUMBER: US/10/015,671A				
; PRIOR FILING DATE: 2001-12-11				
; Prior application removed - See File Wrapper or Palm				
; NUMBER OF SEQ ID NOS: 477				
; SEQ ID NO 275				
; LENGTH: 432				
; TYPE: PRT				
; ORGANISM: Homo sapiens				
US-10-015-671A-275				
Query Match 98.1%; Score 2297.5; DB 2; Length 432;				
Best Local Similarity 98.8%; Pred. No. 4,3e-226;				
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1				
Oy		2	DPDSQPINSLDVKRLRKPRIMETFRKKVGIPIIALSLASTIIVVLIKYILDKRYTL	61
Dd		4	DPDSQPINSLDVKRLRKPRIMETFRKKVGIPIIALSLASTIIVVLIKYILDKRYTL	63
Oy		62	CGQLHFI PRKQLCDGELDCPLGEBDEHCVKSPFGPAVAVALSKDRSTLQVLD SATGNW	121
Dd		64	CGQLHFI PRKQLCDGELDCPLGEBDEHCVKSPFGPAVAVALSKDRSTLQVLD SATGNW	123
Oy		122	FSACPDNTTEALAEFAACROMGYSSAPTPFAVAITEGPDDILDVETITENSOELMRNNSGPC	181
Dd		124	FSACPDNTTEALAEFAACROMGYSSAPTPFAVAITEGPDDILDVETITENSOELMRNNSGPC	178
Oy		182	LGSGLSVSHCLACGSKLTPTRVVGGEEAVSDBMPQVSIQXKHCVCGSLIDPHMVLYLA	241
Dd		179	LGSGLSVSHCLACGSKLTPTRVVGGEEAVSDBMPQVSIQXKHCVCGSLIDPHMVLYLA	238
Oy		242	AHCFRKHADVFNWKVRAGSDKLGSFPPLSAVAKIIIIEFNPMYPKNDIALMLKLOFLTFPS	301
Dd		239	AHCFRKHADVFNWKVRAGSDKLGSFPPLSAVAKIIIIEFNPMYPKNDIALMLKLOFLTFPS	298
Oy		302	GTVRPICLPFDEBELTPATPLMIIGMFTKONGKKMSDILLQASVOVIDSTCNADDAVQ	361
Dd		299	GTVRPICLPFDEBELTPATPLMIIGMFTKONGKKMSDILLQASVOVIDSTCNADDAVQ	358

Oy		362	GEVEEKMKACAIPEGGVDTCCGDSGGPRLMTQSDMHVVGIVSMWGC CGGPSTPGYTRYKS	421
Db		359	GEVEEKMKACAIPEGGVDTCCGDSGGPRLMTQSDMHVVGIVSMWGC GGPGSTPGYTRYKS	418
Oy		422	AYLNIYVMKAEI 435 	
Db		419	AYLNINYNWKAEI 432 	
RESULT 9				
US-10-015-393A-275				
Sequence 275; Application US/10015393A				
Patent No. 6951737				
GENERAL INFORMATION:				
APPLICANT: Baker, Kevin P.				
APPLICANT: Botstein, David				
APPLICANT: Deenoyers, Luc				
APPLICANT: Eaton, Dan I.				
APPLICANT: Ferrara, Napoleone				
APPLICANT: Fong, Sherman				
APPLICANT: Gao, Wei-Qiang				
APPLICANT: Goddard, Audrey				
APPLICANT: Godowski, Paul J.				
APPLICANT: Grimaldi, Christopher J.				
APPLICANT: Gurney, Austin L.				
APPLICANT: Hillan, Kenneth J.				
APPLICANT: Pan, James				
APPLICANT: Paoni, Nicholas F.				
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic				
TITLE OF INVENTION: Acids Encoding the Same				
FILE REFERENCE: P2830PIC46				
CURRENT APPLICATION NUMBER: US/10/015.393A				
CURRENT FILING DATE: 2002-06-10				
Prior Application removed - See File Wrapper or Palm				
NUMBER OF SEQ ID NOS: 477				
SEQ ID NO 275				
LENGTH: 432				
TYPE: PRT				
ORGANISM: Homo sapiens				
US-10-015-393A-275				
Query Match				
Best Local Similarity 98.1%; Score 2297.5; DB 2; Length 432;				
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1				
Oy		2	DPDSDQPLNSIDVPELRKRPIEMETFRKVGIPIIITALLSLASIIIVVLIKVIDKYFL 61	
Db		4	DPDSDQPLNSIDVPELRKRPRIPMETFRKVGIPIIITALLSLASIIIVVLIRKVIDKYFL 63	
Oy		62	CGOPHLFPRKQLDDGLDCPLGDEDEHCVSFPFGPAVNRLSKDRSTLQVLSATGNW 121	
Db		64	CGOPHLFPRKQLCDGLDCPLGDEDEHCVSFPFGPAVNRLSKDRSTLQVLSATGNW 123	
Oy		122	FSACFDNFTEALATACRQMSYSSEKPTFAVEISPDODLDVAEITEINSQELRMNSSGPC 181	
Db		124	FSACFDNFTEALATACRQMGYS-----RAIEIGPDQDLDAVEITEINSQELRMNSSGPC 178	
Oy		182	LSGSLVSIHCTIACGSKIKTTPRVVVGEEBASVDSPWQVSIQYDKOHVCGSILDPHMVTLA 241	
Db		179	LSGSLVSIHCTIACGSKIKTTPRVVVGEEBASVDSPWQVSIQYDKOHVCGSILDPHMVTLA 238	
Oy		242	AHCRKHHTDVNMKVVRAGSDLGSPSLAAVKIIITERNPMYPKDNDIALMKLQPLTFPS 301	
Db		239	AHCRKHHTDVNMKVVRAGSDLGSPSLAAVKIIITERNPMYPKDNDIALMKLQPLTFPS 298	
Oy		302	GTVVPICLPFDEBELTATPLMIIGWGFTKONGGMSDILLQASQVUIDSTRCANADAYQ 361	
Db		299	GTVVPICLPFDEBELTATPLMIIGWGFTKONGGMSDILLQASQVUIDSTRCANADAYQ 358	
Oy		362	GEVEEKMKACAIPEGGVDTCCGDSGGPRLMTQSDMHVVGIVSMWGC GGPGSTPGYTRYKS 421	
Db		359	GEVEEKMKACAIPEGGVDTCCGDSGGPRLMTQSDMHVVGIVSMWGC GGPGSTPGYTRYKS 418	

```
Qy 182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
|
|
|
Db 179 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238
|
|
|
Qy 242 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 301
|
|
|
Db 239 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 298
|
|
|
Qy 302 GTVRPICLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVQVIDSTRCANADAYQ 361
|
|
|
Db 299 GTVRPICLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVQVIDSTRCANADAYQ 358
|
|
|
Qy 362 GEVTERKMCAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 421
|
|
|
Db 359 GEVTERKMCAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 418
|
|
|
Qy 422 AYLNWYVWKAEL 435
|
|
|
Db 419 AYLNWYVWKAEL 432
|
|
|
```

RESULT 6

```
US-10-015-389A-275
; Sequence 275, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; PRIORITY FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-015-389A-275
```

```
Query Match 98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
Qy 2 DPDSQDPLNSLDVPLKRPRIIPMETFRKVGIPITIALSLASIIIVVLLKVIIDKYFL 61
|
|
|
Db 4 DPDSQDPLNSLDVPLKRPRIIPMETFRKVGIPITIALSLASIIIVVLLKVIIDKYFL 63
|
|
|
Qy 62 CGQPLHFIIPKQCLDGEIDCPLEGDEBHCYKSPFEGPAVAVRLSKDSTLOVLSATGNW 121
|
|
|
Db 64 CGQPLHFIIPKQCLDGEIDCPLEGDEBHCYKSPFEGPAVAVRLSKDSTLOVLSATGNW 123
|
|
|
Qy 122 FSAFDFNTEALATACRQWYSKPTFRFAVEIGPDODLVVEITENSQELRMNNSGPC 181
|
|
|
Db 124 FSAFDFNTEALATACRQWYSKPTFRFAVEIGPDODLVVEITENSQELRMNNSGPC 178
|
|
|
Qy 182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
|
|
|
Db 179 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238
|
|
|
```

```
Qy 242 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 301
|
|
|
Db 239 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 298
|
|
|
Qy 302 GTVRPICLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVQVIDSTRCANADAYQ 361
|
|
|
Db 299 GTVRPICLPFDEBELTPATPLMIIGMFTKONGKMSDILLQASVQVIDSTRCANADAYQ 358
|
|
|
Qy 362 GEVTERKMCAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 421
|
|
|
Db 359 GEVTERKMCAGIPBGGVDTCCGDSGGPLMYQSDQMHVGVISWGYCGGSPSTPGVYTKVS 418
|
|
|
Qy 422 AYLNWYVWKAEL 435
|
|
|
Db 419 AYLNWYVWKAEL 432
|
|
|
```

RESULT 7

```
US-10-006-768A-275
; Sequence 275, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Deenoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006,768A
; PRIORITY FILING DATE: 2002-03-05
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-768A-275
```

```
Query Match 98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

```
Qy 2 DPDSQDPLNSLDVPLKRPRIIPMETFRKVGIPITIALSLASIIIVVLLKVIIDKYFL 61
|
|
|
Db 4 DPDSQDPLNSLDVPLKRPRIIPMETFRKVGIPITIALSLASIIIVVLLKVIIDKYFL 63
|
|
|
Qy 62 CGQPLHFIIPKQCLDGEIDCPLEGDEBHCYKSPFEGPAVAVRLSKDSTLOVLSATGNW 121
|
|
|
Db 64 CGQPLHFIIPKQCLDGEIDCPLEGDEBHCYKSPFEGPAVAVRLSKDSTLOVLSATGNW 123
|
|
|
Qy 122 FSAFDFNTEALATACRQWYSKPTFRFAVEIGPDODLVVEITENSQELRMNNSGPC 181
|
|
|
Db 124 FSAFDFNTEALATACRQWYSKPTFRFAVEIGPDODLVVEITENSQELRMNNSGPC 178
|
|
|
Qy 182 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 241
|
|
|
Db 179 LSGSLVSLHCLACGKSLKTPRVVGGEEASVDSWPMQVSIQYDKQHVCGGSLIDPHWVLT 238
|
|
|
Qy 242 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 301
|
|
|
Db 239 AHCRKHTDVFNWVKVRAGSDKLSFPSLAVALKIIIEFNPMYPKNDIALMKLOFPLTFS 298
|
|
|
```



```
Query Match          99.8%; Score 2338; DB 2; Length 435;
Best Local Similarity 99.8%; Pred. No. 3.1e-230;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSQPLNSLDVKKLRKPRIPMETFRKVGIPPIIALSLASIIIVVLIKYLIDKXYFL 60
DB 1 MDPDSQPLNSLDVKKLRKPRIPMETFRKVGIPPIIALSLASIIIVVLIKYLIDKXYFL 60
QY 61 LCGQPLHFIIPRKOLCGELDCPLGEBDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
DB 61 LCGQPLHFIIPRKOLCGELDCPLGEBDEHCVKSPFEGPAVAVRLSKDRSTLQVLD SATGN 120
QY 121 WFSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 180
DB 121 WFSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 180
QY 121 WFSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 180
DB 121 WFSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 180
QY 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCCGSLIDPHWVLT 240
DB 181 CLSGSLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCCGSLIDPHWVLT 240
QY 241 AAHCFRKHDTDFVEMKVRASGDKLGSFPSLAIVAKIIIEENPMYPKNDIALMKLQPLTF 300
DB 241 AAHCFRKHDTDFVEMKVRASGDKLGSFPSLAIVAKIIIEENPMYPKNDIALMKLQPLTF 300
QY 301 SGTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
DB 301 SGTVRPCLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 360
QY 361 QSEVTERKMCAGIPFEGGVDTCCGDSGGLMYOSDQMHVVGIVSMGCGGSPSTPGVYTV 420
DB 361 QSEVTERKMCAGIPFEGGVDTCCGDSGGLMYOSDQMHVVGIVSMGCGGSPSTPGVYTV 420
QY 421 SAYLNIYVWKRAEL 435
DB 421 SAYLNIYVWKRAEL 435

RESULT 4
US-09-851-588-8
; Sequence 8, Application US/09851588
; Patent No. 6682890
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C.
; APPLICANT: Wilson, Keith E.
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
; FILE REFERENCE: A-68829-1/DJL/JJD/AMS
; CURRENT APPLICATION NUMBER: US/09/851,588
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/642,252
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: US 09/656,002
; PRIOR FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-851-588-8

Query Match          99.8%; Score 2337; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 4e-230;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 122 FSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 181
DB 124 FSACDNTEALAEACRQMGYSKPTFRVAIEIGPDOLDVVEITENSQELMRNSSGPG 183
QY 182 LSGSLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCCGSLIDPHWVLT 241
DB 184 LSGSLVSLHCLACGSLKTPRVVGGEEASVDSMPQVSIQYDKQVCCGSLIDPHWVLT 243
QY 242 AAHCFRKHDTDFVEMKVRASGDKLGSFPSLAIVAKIIIEENPMYPKNDIALMKLQPLTF 301
DB 244 AAHCFRKHDTDFVEMKVRASGDKLGSFPSLAIVAKIIIEENPMYPKNDIALMKLQPLTF 303
QY 302 GTVRPICLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 361
DB 304 GTVRPICLPFPDEELTPATPLMIIGWFTKQNGKMSDILLQASVQVIDSTRCNADAY 363
QY 362 GEVTERKMCAGIPFEGGVDTCCGDSGGLMYOSDQMHVVGIVSMGCGGSPSTPGVYTV 421
DB 364 GEVTERKMCAGIPFEGGVDTCCGDSGGLMYOSDQMHVVGIVSMGCGGSPSTPGVYTV 423
QY 422 AYLNIYVWKRAEL 435
DB 424 AYLNIYVWKRAEL 437

RESULT 5
US-10-012-231A-275
; Sequence 275, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 275
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-231A-275

Query Match          98.1%; Score 2297.5; DB 2; Length 432;
Best Local Similarity 98.8%; Pred. No. 4.3e-226;
Matches 429; Conservative 0; Mismatches 0; Indels 5; Gaps 1;
```

Db 361 QGEVTERKMCAGIPGEGVDTCQDGS GGPLMYQSDQMHWVGIVSWGCGGSPTPGVYTKV 420
QY 421 SAYLNIWYWKAEI 435
Db 421 SAYLNIWYWKAEI 435

RESULT 2
US-09-008-271A-6

; Sequence 6, Application US/09008271A
; Patent No. 6203979
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/008,271A
FILING DATE: 16-Jan-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 6 :

US-09-008-271A-6

Query Match 99.8% Score 2338; DB 2; Length 435;
Best Local Similarity 99.8%; Pred. No. 3,1e-230;
Matches 434; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDPDSDDPLNSIDVYPLAKPRIPMETFRKVGPIIIIALSLASIIIVVVLKVIIDKYYF 60
Db 1 MDPDSDDPLNSIDVYPLAKPRIPMETFRKVGPIIIIALSLASIIIVVVLKVIIDKYYF 60
QY 61 LCGGPIAHFIPKQOLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTIQVLD SATGN 120
Db 61 LCGGPIAHFIPKQOLCDGELDCPLGDEDEHCVKSPFEGPAVAVRLSKDRSTIQVLD SATGN 120
QY 121 WFSACFDNFTEALATACRQNGYSSKPTFRFAVEIGPDODLDVVEITENSQELRMNNSGP 180
Db 121 WFSACFDNFTEALATACRQNGYSSKPTFRFAVEIGPDODLDVVEITENSQELRMNNSGP 180

QY 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSPWQVSIQYDKQHYCGGSIIDPHWVLT 240
Db 181 CLSGSLVSLHCLACGKSLKTPRVVGGEEASVDSPWQVSIQYDKQHYCGGSIIDPHWVLT 240
QY 241 AAHCFRRKTDVFNKVRAGSDKLGSPPSLAVAKIIIEFNPMYRKNDIALMKIQPLTF 300
Db 241 AAHCFRRKTDVFNKVRAGSDKLGSPPSLAVAKIIIEFNPMYRKNDIALMKIQPLTF 300
QY 301 SGTVRPICLPPEDELTPATPMTIGWFTKONGKXSDIILQASVQVISTRCNADAY 360
Db 301 SGTVRPICLPPEDELTPATPMTIGWFTKONGKXSDIILQASVQVISTRCNADAY 360
QY 361 QGEVTERKMCAGIPGEGVDTCQDGS GGPLMYQSDQMHWVGIVSWGCGGSPTPGVYTKV 420
Db 361 QGEVTERKMCAGIPGEGVDTCQDGS GGPLMYQSDQMHWVGIVSWGCGGSPTPGVYTKV 420
QY 421 SAYLNIWYWKAEI 435
Db 421 SAYLNIWYWKAEI 435

RESULT 3
US-09-968-415-6

; Sequence 6, Application US/09968415
; Patent No. 6855811
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi

TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/968,415
FILING DATE: 26-Sep-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/659,151
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 435 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 1337018
SEQUENCE DESCRIPTION: SEQ ID NO: 6 :

US-09-968-415-6

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 5, 2005, 13:45:44 ; Search time 45 Seconds

(Without alignments)
799.198 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342

Sequence: 1 MDPDSQPLNSLDVFKLRKP.....VYTKSAVIMWYVWRAEL 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/iaa/5.COMB.pep.*
2: /cgn2_6/prodata/1/iaa/6.COMB.pep.*
3: /cgn2_6/prodata/1/iaa/H.COMB.pep.*
4: /cgn2_6/prodata/1/iaa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/iaa/R.COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2342	100.0	435	2	US-09-607-745-2
2	2338	99.8	435	2	US-09-008-271A-6
3	2338	99.8	435	2	US-09-868-415-6
4	2337	99.8	437	2	US-09-851-588-8
5	2297.5	98.1	432	2	US-10-012-231A-275
6	2297.5	98.1	432	2	US-10-015-389A-275
7	2297.5	98.1	432	2	US-10-006-768A-275
8	2297.5	98.1	432	2	US-10-015-671A-275
9	2297.5	98.1	432	2	US-10-015-393A-275
10	2297.5	98.1	432	2	US-10-011-833A-275
11	2297.5	98.1	432	2	US-10-006-041A-275
12	2297.5	98.1	432	2	US-10-012-064A-275
13	2196	93.4	423	2	US-09-656-002-2
14	2188	93.4	406	2	US-09-851-588-6
15	1266	54.1	292	2	US-09-607-745-9
16	777	33.2	418	2	US-10-177-661-6
17	698.5	29.8	477	2	US-10-177-661-2
18	697.5	29.8	562	2	US-09-879-792-12
19	683	29.2	446	2	US-10-177-661-4
20	677.5	28.9	492	2	US-09-885-166A-895
21	677.5	28.9	492	2	US-09-879-792-14
22	677.5	28.9	492	2	US-09-679-426-895
23	677.5	28.9	492	2	US-09-759-143-895
24	677.5	28.9	492	2	US-10-012-896-895
25	676.5	28.9	492	2	US-09-342-749-2
26	676.5	28.9	492	2	US-09-691-840-2
27	676.5	28.9	510	2	US-09-949-016-11074

28	675.5	28.8	393	2	US-09-759-143-934	Sequence 934, App
29	675.5	28.8	393	2	US-10-012-896-934	Sequence 934, App
30	675.5	28.8	492	2	US-09-759-143-932	Sequence 932, App
31	675.5	28.8	492	2	US-10-012-896-932	Sequence 932, App
32	660.5	28.2	521	2	US-09-949-016-11081	Sequence 11081, A
33	660.5	28.2	521	2	US-09-949-016-11082	Sequence 11082, A
34	660.5	28.2	521	2	US-09-949-016-11083	Sequence 11083, A
35	658	28.1	453	2	US-09-999-833A-69	Sequence 69, App1
36	658	28.1	453	2	US-10-020-445A-69	Sequence 69, App1
37	655.5	28.0	454	2	US-09-518-046-2	Sequence 2, App1
38	655.5	28.0	454	2	US-09-650-371-2	Sequence 2, App1
39	628.5	26.8	445	2	US-09-856-371A-8	Sequence 8, App1
40	588.5	25.1	417	2	US-09-820-002-4	Sequence 4, App1
41	584.5	25.0	452	2	US-09-949-016-7182	Sequence 7182, Ap
42	584	24.9	455	2	US-09-261-416-2	Sequence 2, App1
43	580	24.8	376	2	US-09-820-002-2	Sequence 2, App1
44	574	24.5	416	1	US-09-000-846-2	Sequence 2, App1
45	574	24.5	457	2	US-09-856-371A-10	Sequence 10, App1

ALIGNMENTS

RESULT 1
US-09-607-745-2
; Sequence 2, Application US/09607745
; Patent No. 6750034
; GENERAL INFORMATION:
; APPLICANT: Dartow, Andrew L
; APPLICANT: Qi, Jai-shen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: DNA encoding human serine protease D-G
; FILE REFERENCE: ORT-1273
; CURRENT APPLICATION NUMBER: US/09/607,745
; CURRENT FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-607-745-2

Query Match	Best Local Similarity	100.0%; Score 2342; DB 2; Length 435;
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MDPDSQPLNSLDVFKLRKPIMETFRKVGIPITIIALLSLASIIIVVLIKVILDKYF 60
DB	1	MDPDSQPLNSLDVFKLRKPIMETFRKVGIPITIIALLSLASIIIVVLIKVILDKYF 60
QY	61	LCGQPLHPIPRKOLCDGELDCPLGDEDEHCYKSPPEGAVAVRASKORSTIQVDSATGN 120
DB	61	LCGQPLHPIPRKOLCDGELDCPLGDEDEHCYKSPPEGAVAVRASKORSTIQVDSATGN 120
QY	121	WFSACFDNFTEALAEATACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP 180
DB	121	WFSACFDNFTEALAEATACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP 180
QY	121	WFSACFDNFTEALAEATACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP 180
DB	121	WFSACFDNFTEALAEATACROWGYSKPTFAVEIGPDODLVEITENSQELMRNNSGP 180
QY	181	CLSGLVSLHCLACGKSLKTRPVVVGEBASVDSWPMOVSIOYDKQHYCGGSIIDPHVLT 240
DB	181	CLSGLVSLHCLACGKSLKTRPVVVGEBASVDSWPMOVSIOYDKQHYCGGSIIDPHVLT 240
QY	241	AAHCRKHTDVFNKVRAGSDKLSPSLIAVAKITIIIEPNMVPKNDIALMKIOPLTTF 300
DB	241	AAHCRKHTDVFNKVRAGSDKLSPSLIAVAKITIIIEPNMVPKNDIALMKIOPLTTF 300
QY	301	SGTVRPICLPFDEBELPATPLWTIIGGFTKONGKMSDILLQASVOVIDSTRCNADAY 360
DB	301	SGTVRPICLPFDEBELPATPLWTIIGGFTKONGKMSDILLQASVOVIDSTRCNADAY 360
QY	361	QGVTEKMKMGITIEGGVDTQGSGLPMTQSDQMHVGVISVKGCCGPGSTGVYTKV 420

THIS PAGE BLANK (USPTO)

QY	402	ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrIysValSer	421
Db	1484	GTTAGCTGGGGCTATGGCTGCGGGGGCCGAGCACCCCAAGATATACACCAAGGCTCTCA	1543
QY	422	AlaTyrLeuAsnTrpIleTyrAsnValTrpIleAlaGluLeu	435
Db	1544	GCCTATCTCACTGATGATCTTACAAATGTCGTGAAAGGCTGAGCTG	1585

Search completed: December 6, 2005, 20:22:41
Job time : 710 secs

PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamathkar S, Kovatis SG;
PI Meyers RE, Morrissey WF, Olandt FJ, Sen A, Viedy PO, Mills GB;
PI Baet RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX
DR MPI; 2002-723277/78.
DR P-PSDB; ABG96430.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
PS Disclosure: Page 438; 481pp; English.
XX
CC The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterizing cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. non-tuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present nucleic acid
CC sequence encodes one of the ovarian cancer markers described in the
CC invention
XX
SQ Sequence 2307 BP; 555 A; 647 C; 614 G; 460 T; 0 U; 31 Other;
Alignment Scores:
Pred. No.: 6,95e-207 Length: 2307
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0
US-10-803-530-2 (1-435) x ABS76529 (1-2307)
QY 2 ASPPCASPSPASGGLPRLQENLSENLEAAPPVALVLPRLQENLRGYPRLQY 21
DB 24 GATCTCTGACAGTGAACCTCTGACACCTCTGATGATCAACCCCTGGCAACCCCGT 343
QY 22 ILEPROMETGLUTHPHEARGLYSVALGLYLEPRLQELLEILEALALEUENSERLEU 41

DB 344 ATCCCATGAGACCTTCAGAAAGTGGAGATCCCATCATCATAGACCTG 403
QY 42 ALASERLEILEILEVALVALLLEUULEYSVALILEUASPLYSYRTRYRPHLEU 61
DB 404 GCGAGTATCATATGAGTGTGCTCATAGAGTGAATTCGATTAATTAATCTTCTC 463
QY 62 CYSGLYLINPRLQENHISPHLEILEPRLQYLSGLINLEUCYASBPGLYLULEUASP 81
DB 464 TGGGGGAGGCTCTCCATCTTCATCCGAGAGAGAGCTGTGTGACGAGAGCTGACTGT 523
QY 82 PRLQENGLYLUASBPGLUHLHISCYEVALYLSERPHPRQGLUGLYPRLQALVALLA 101
DB 524 CCTTGGGGGAGGAGCAGAGAGCACTGTGTCAAGACTTCCCGGAGGCTTGCAGTGGCA 583
QY 102 VALRGLSEUERYLSAPASERPTHLEUGLVALLEUASPRALATPRLQYASRTP 121
DB 584 GTCCGCTCTCAAGAGACCATGACACTGAGAGTGTGACTGGCCACAGGGAACCTGG 643
QY 122 PHESERLACYSPPHEAPSPNPHETHRGLVALALEUVALGLUTRALACYSARGINMET 141
DB 644 TTCTCTGCTGTTTCGACACTTCACAGAGCTTCGCTGACAGAGCTGTAGGACAGATG 703
QY 142 GLYTYRSERSELYSPROTHRPHARGALAYALGLULEGLYPROASBPGLINPRLQEN 161
DB 704 GGCTACAGACAGCAAAACCATTTTCAGAGCTGTGAGAGATTGGCCGACAGATCTGAT 763
QY 162 VALVALGLULETHRGLUASERGLINLEUASRGLINLEUASRSEGLYPRQCY 181
DB 764 GTTGTGAATACAGAAAGAACGACAGAGCTTGCATGCGAACTCAAGTGGCCCTGT 823
QY 182 LEUSERGLSERLEUVALSERLEUHLISCYSEUVALACYSGLYLSERLEUPLYTHRPRO 201
DB 824 CTCCTCAGGCTCCCTGCTGCTCTCCCTGACTGTCTTCCTGTGGAGAGAGCTGAACCC 883
QY 202 ARGVALVALGLYGLUGLUALASERVALASBPSETRTPROTPGLINVALSERLEGLIN 221
DB 884 CGTGTGTGGGTGGGAGGAGGCTCTGTGATTTCTTGCTGGCAGTGCAGATCCAG 943
QY 222 TYRAPHLYSGNHLSEVALCYSGLYGLYSERLEULEUASPRPHISTPVALLEUTHALA 241
DB 944 TACACAAACGACGCTGTGGAGGAGCATCTGACCCCACTGGGTCTCAACGGCA 1003
QY 242 ALAHISCYSPHEARGLYSHISTHASPVALPHEASNTPLYSVALARGALAGLYSERASP 261
DB 1004 GCCACCTGCTTCAGAAACATACCGATGTGTCAACTGGAAGTCCGGCAGGCTCAGAC 1063
QY 262 LYSLEUGLYSERPHPRQSELEUVALVALALYSILEILEILEGLUPHEASNP 281
DB 1064 AAACCTGGGACGCTCCCATCTGCTGTGGCCAAAGATCATCATTAATTCAAACCCC 1123
QY 282 METTYRPROLYSAPASNPSPHLEALEUWETLYSEUGLINPHEPRLQENLTHRPHSE 301
DB 1124 ATGTACCCCAAGCAATGACATGCTCAATGAAAGCTGATTCACATCTTCTCA 1183
QY 302 GLYTHRVALARGPROILECYSELEUPROPHESAPGLUGLULEUTHRPROALATHRPRO 321
DB 1184 GGCACAGTCAAGCCCATCTGCTGCTCTTTATATGAGAGCTCATCTCCAGCCACCCA 1243
QY 322 LEUTRPILEILEGLYTPGLYPHETHRILYSGLINASNGLYLYSMESERASPILEU 341
DB 1244 CTCTGATCATTTGATGGGGCTTTACGAAGACAGATGAGGGAAGATGTCTGACATATG 1303
QY 342 LEUGLINALASERVALGINVALILEASPSERTHRARGYSAENALASPSAPALATYCT 361
DB 1304 CTGCAAGGCTCAGTCCAGTCAATTCAGACGACACAGCTGCAATGACAGATGCTGAC 1363
QY 362 GLYGLUVALTRGULYSMETMETCYEALAGLYLEPRLQGLUGLYVALASPTHRCYS 381
DB 1364 GGGGAAGTCAACGAGAAATATATGTGTGACAGCATCCCGAAGGGGTGTGACACCTGC 1423
QY 382 GINGLYASBPSEGLYGLYPROLEUMETRYGLINSEASBPGLINTPRLPHSVALVALL 401
DB 1424 CAGGTGACAGTGTGGGCTCCCTGATGTACCAATCTGACAGTGGCATGTGTGGGCATC 1483

XX Ashkenazi A, Goddard A, Gurney A, Polakis P, Smith V, Wood WI;
PI Wu T, Zhang Z;
XX WPI; 2005-418022/42.
DR P-PSDB; AEA54849.
XX
PT New polynucleotides encoding tumor-associated antigenic target (TAT)
PT polypeptides, useful e.g., in gene therapy, in diagnosis of cancer, and
PT in the development of cancer therapeutics.
XX
PS Claim 2; SEQ ID NO 2; 137bp; English.
XX
CC The invention relates to novel tumor associated antigenic target (TAT)
CC polypeptides and nucleic acid molecules encoding such polypeptides.
CC Nucleic acid molecules of the invention are useful as hybridization
CC probes in chromosome and gene mapping, for producing TAT polypeptides, in
CC gene therapy for treating TAT expressing tumors and in the diagnosis of
CC such tumors. They are also useful for producing transgenic animals for
CC screening and developing therapeutically useful reagents, in tissue
CC typing and for producing probes for PCR, Northern analysis, Southern
CC analysis and Western analysis. The present sequence is the human TAT502
XX cDNA.
SQ Sequence 2104 BP; 499 A; 604 C; 577 G; 424 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6.15e-207 Length: 2104
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 14 Gaps: 0
US-10-803-530-2 (1-435) x AEA54844 (1-2104)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuAlaGlySerProAla 21
Db 242 GATCTGACGATGATCACTTCAACAGCCCTGATGTCAACCCCTGCGCAAAACCCCGT 301
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 302 ATCCCAATGAGACCTTCAGAAAGGTGGGATCCCATCATCTATACATCTACAGACCTG 361
QY 42 AlaserIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
Db 362 GCGAGTATCATCTATTGGTTGTTCTTCATCAAGGTATCTGGATTAATACTACTTCCTGC 421
QY 62 CysGlyGlnProLeuAsnIlePheIleProArgLysGlnLeuCysAspGlyValLeuAspCys 81
Db 422 TGGGGGAGGCGCTCTCCACTTCATCCGAGGAGACACTGTGTGACGAGAGCTGGACTGT 481
QY 82 ProLeuGlyValLeuAspGluGlnIleCysValIysSerPheProGluGlyProAlaValAla 101
Db 482 CCTTGGGGAGGAGCAGGAGGAGCACTGTGTCAAGAGCTTCCCGAAGGGGCTGCACTGCA 541
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleAspSerAlaThrGlyAsnTrp 121
Db 542 GTCCGCGCTCTCCAGAGACCGATCCACATGCAAGGTGCTGACTCCGCCACAGGGAACTGG 601
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 602 TTCTCTGCTGCTTTCGACAACTTCACAGAGCTCTCCGTGAGACGCTGTAGGCAATG 661
QY 142 GlyTyrSerSerLysProThrPheArgLysValGlyIleGlyProAspGlnAspLeuAsp 161
Db 662 GCGTACAGAGCAAAACCACTTTCAGACTGTGAGATGGCCCAAGACAGACTCTGAT 721
QY 162 ValValGluIleThrGlnAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 722 GTTGTGAATCACAGAAACAGCCAGAGAGCTTGCAATGCGAACTCAAGGGGCGCTGT 781
QY 182 LeuSerGlySerLeuValSerLeuAsnIleCysLeuAlaCysGlyLysSerLeuLysThrPro 201

Db 782 CTCTCAGAGCTCCCTGGATCTCCCTGCACCTGTCTTGCTGTGGGAAGAGCTGAAGACCC 841
QY 202 ArgValValGlyValGluGlnAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 842 CTTGTGGTGGTGGGAGAGAGGCTCTGTGTGATTTCTTGCCCTTGGCAGTGACATCCAG 901
QY 222 TyrAspLysGlnIleValCysGlyLysSerIleLeuAspProHisTrpValLeuThrAla 241
Db 902 TACGACAAACAGACAGCTCTGTGAGAGGACATCTTGAGCCCACTGGAGTCTTCAAGGCA 961
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 962 GCCCACTGCTTCAGAAACATACCGATGTGTCACTGAAGGTGGCGAGCTCAAGAC 1021
QY 262 LysLeuGlySerPheProSerIleuValAlaLysIleIleIleIleGluPheAsnPro 281
Db 1022 AACTGGGCGAGTTCCTCCATCCCTGCTGTGGCCAAAGTCAATCATCATTAATTCACCCC 1081
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACATGACATGACCTCCCTCATGAGCTGCAGTTCCACTCATTTCTCA 1141
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGlnLeuThrProAlaThrPro 321
Db 1142 GGCACAGTCAGGCCCATCTGTCTGCTCTTGTGATGAGAGCTCACTCCAGCAACCCCA 1201
QY 322 LeuTrpIleIleGlyTyrPglPheThrLysGlnAsnGlyValLysMetSerAspIleLeu 341
Db 1202 CTCTGGATCATTTGATGGGCTTTTACAGACAGAAATGAGAGAAATGCTGTGACATACAG 1261
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1262 CTGAGGGGTCAAGCCAGCTCATTTGACAGACACGGGTGCAATGCAATGATGTCACAG 1321
QY 362 GlyGluValIleThrGluMetMetCysAlaGlyIleProGluGlyValIleAspThrCys 381
Db 1322 GGGGAAATCACCGAAAGATGATGTGACAGGATCCCGAAGGGGTGTGACACTCTGC 1381
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1382 CAGGTGACAGTGTGGGCGCCCTGATGTCATCTGACCAATGGATGTGTGGGCAATC 1441
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1442 GTTAGCTGGGGCTATGCTGCGGGGGCCCGAGACCCAGAGATATACCAAGCTCTCA 1501
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValIleTyrLysAlaGluLeu 435
Db 1502 GCTATCTCAACTGATCTTCAATGTCTGGAAGGCTGAGCTG 1543
RESULT 15
ID ABS76529 standard; cDNA; 2307 BP.
XX
AC ABS76529;
XX
DT 11-DEC-2002 (first entry)
XX
DE cDNA encoding human ovarian cancer marker OV86.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW non tuberculous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.

by OSNA, an antibody or its fragment that specifically binds to the polypeptide, determining or monitoring the presence of an ovarian specific protein in a sample, diagnosing or monitoring the presence and metastases of ovarian cancer in a patient, a kit for detecting a risk of cancer or presence of cancer in a patient (comprising a means for determining the presence of OSNA or the polypeptide), treating a patient with ovarian cancer and a vaccine comprising the polypeptide or the nucleic acid encoding the polypeptide. The sequences, kits and methods are useful for diagnosing or monitoring the presence and metastases of ovarian cancer in a patient and for detecting a risk of cancer or presence of cancer in a patient, as well as for staging and imaging ovarian cancer. The sequences, vaccine, composition and method are useful for treating a patient with ovarian cancer. The invention discloses two sets of OSNA (cDNAs and their splice variants), Ovr15, the gene encoding which is located on chromosome 11q23.3 and Pro104, the gene encoding which is located on chromosome 16p13.3. The present sequence is the Ovr15 cDNA.

XX Sequence 2104 BP; 499 A; 604 C; 577 G; 424 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,15e-207 Length: 2104
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 14 Gaps: 0

US-10-803-530-2 (1-435) x ABW04403 (1-2104)

OY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspVallysProLeuArgLysProArg 21
DB 242 GATCCTGACAGTGATCAACCTCTGAAACGCTCGATGCAAAACCTTCGGCAAAACCCCGT 301
OY 22 IleProMetGlnThrPheArgLysValGlyIleProIleIleIleIleLeuLeuSerLeu 41
DB 302 ATCCCATGAGACCTTCAGAAAGGTGGGAGATCCCATCATCATATGACATCTAGCTGAGCTG 361
OY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
DB 362 GCGAGTATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 421
OY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCyAspGlyGlyLeuAspCys 81
DB 422 TGGGGGAGCCTCTCATTATCCCGAAGAGAGCTGTGTGAGAGAGAGCTGAGCTGAGCTG 481
OY 82 ProLeuGlyGlyAspGlyGlnHisCysVallysSerPheProGlyIleProAlaValAla 101
DB 482 CCTTGGGGAG 541
OY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
DB 542 GTCCCTCTCTCCAAAGACCGATCCACCTGAGGAGCTGTGATCTGCGCCACAGAGAACTGG 601
OY 122 PheSerAlaCySPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCySPheArgGlnMet 141
DB 602 TTCTCTGCTGCTTTGACACATCTTCAAGAGCTCTCGCTGAGACAGCCTCTGAGAGAGTGG 661
OY 142 GlyIleSerSerLysProThrPheArgAlaValGlyIleGlyProAspGlnAspLeuAsp 161
DB 662 GGCTACACAGAGAAACCACTTTCAGAGCTGTGAGATGGCCAGACAGAGATCTGAGAT 721
OY 162 ValValGlyIleThrGlnAsnSerGlnIleuAspMetArgAsnSerSerGlyProCys 181
DB 722 GTTGTGTAATACAGAAACAGCCAGAGAGCTTCCATGCGGAACTCAAGTGGGCGCTGT 781
OY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCySPheGlyLysSerLeuLysThrPro 201
DB 782 CTCTCAGGCTCCCTGCTCTCTGCACTGTCTTGGCTGTGGGAAGAAGCTGAAGACCCG 841
OY 202 ArgValIleGlyGlyGlnIleuAspSerValAspSerTrpProTrpGlnValSerIleGln 221
DB 842 CGTGTGTGTGGTGGGAGAGAGAGCTCTGTGATCTTGTGCGCTTGGACAGTCAAGTCCAG 901

OY 222 TyrAspLysGlnHisValCySPheGlyIleSerIleLeuAspProHisTrpValIleuThrAla 241
DB 902 TAGACAAACAGCAGCTGTGTGAGAGAGATCTTGAGCCCACTGGGTCTTCACGGCA 961
OY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValAlaGlyIleSerAsp 261
DB 962 GCCCAGCTCTTCAGAAACATACGATGTGTTCAACTGGAAGGTGGCGGCGGTCTCAGC 1021
OY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAspPro 281
DB 1022 AAATGGGACCTTCCATCTCTGCTGTGTGAGAGAGATCATCATCATATTAATTCACCC 1081
OY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1082 ATGATCCCAAAAGCAATGATCATGACCTTATGAAGGTGAGAGTCCCACTCATCTTCTCA 1141
OY 302 GlyThrValArgProIleCysLeuProPhePheAspGlyGlnLeuThrProAlaThrPro 321
DB 1142 GGCACAGTCAGGCCCATCTGTCTGCTCTTCTTATGATGAGAGCTCACCTCCAGCACCCCA 1201
OY 322 LeuTrpIleIleGlyIleTrpGlyPheThrLysGlnAsnGlyIleLysMetSerAspIleLeu 341
DB 1202 CTCTGATCATTTGATGGGCTTTTACAGAGAGATGAGAGAGATGTGACATTAATCTG 1261
OY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1262 CTGCAAGGCTGATGCTCAGATCATTTGACAGCAGCAGGTGTGATGACAGAGATCCGTA 1321
OY 362 GlyGlnValThrGlyLysMetMetCysAlaGlyIleProGlyGlyIleValAspThrCys 381
DB 1322 GGGGAGTCAACCGAAGAGATGATGTGACAGCATCCCGAAGGGGTGTGAGACCTGC 1381
OY 382 GlnGlyAspSerGlyIleProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
DB 1382 CAGGTGACAGTGTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1441
OY 402 ValSerTrpGlyIleGlyCysGlyIleProSerThrProGlyValIleThrLysValSer 421
DB 1442 GTTACCTGGGCTATGCTGTGGGGGCGCCAGAGCAGCAGGAGATATACCAAGTCTCA 1501
OY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 435
DB 1502 GCTATCTCACTGATCTCAATGTCTGAAAGGCTGAGCTG 1543

RESULT 14
AEA54844
ID AEA54844 standard; cDNA; 2104 BP.
AC AEA54844;
XX 11-AUG-2005 (first entry)
DT
DE Human TAT502 cDNA clone DNA279661, SEQ ID NO: 2.
KW transgenic animal; mapping; gene therapy; tissue typing; tumor;
KW cytosolic; neoplasm; antigen; TAT502; gene; ss.
XX Homo sapiens.
OS
XX
XX
FH Key Location/Qualifiers
FT CDS 233..1546
FT /tag= a
FT /product= "Human TAT502 protein"
FN
PN WO2005052191-A2.
XX
XX 09-JUN-2005.
PD
XX
PF 17-NOV-2004; 2004WO-US038669.
XX
XX 20-NOV-2003; 2003US-0523856P.
PR
XX
PA (GETH) GENENTECH INC.

```

Db      302 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGACATCTGAGCCTG 361
Qy      42 AlaserllellelvalvalleuileysvalilleuapblystrlyrPhleu 61
Db      362 GCGAGTATCATATTGAGTGTCTCATAGAGTATTTGGAATTAATCTACTTCC 421
Qy      62 CysGlyGlnProleuHlspeHleProArgLysGlnLeuCyaaSpGlyLuleuapCyS 81
Db      422 TGGGGGAGCCTCTCCACTTCATCCCGAGAACAGCTGTGTACCGAGAGGCTGACTGT 481
Qy      82 ProleuGlyLuleuapGlyLuleuHlsCysValLysSerPhleProGlyLuleu 101
Db      482 CCTTGGGGGAGGAGCAGAGGACATGTGTCAAGAGCTTCCCGAAGGCTCAGTGGCA 541
Qy      102 ValArgLeuSerLysaPaSerThrLeuGlnValLeuaspSeralanthGlyasnrp 121
Db      542 GTCCGCTCTCCAAAGACCGATCCACATCGAGGTCTGTGATCCGCCACAGGGAACTGG 601
Qy      122 PheSerAlaCysPheaspbnPheHrGlnAlaLeuAlaGlyThrAlaCysArgGlnMet 141
Db      602 TTCTGTGCTGTTCACAACTTCACAGAGCTCTGTGAGACAGCTGTAGGCAGATG 661
Qy      142 GlyTyrSerSerLysProThrPheArgAlaValGlyLileGlyProaspGlyLuleuap 161
Db      662 GGCTACAGCAGCAAACTTCAGAGCTGTGAGATTGGCCACAGACAGGATCTGAT 721
Qy      162 ValValGlyLileThrGlnaspSerGlnLuleuArgMetArgasnsSerSerGlyProCys 181
Db      722 GTTGTGAAATACAGAAACAGCCAGAGGCTTCCCATCGGAACTCAAGTGGGCTGT 781
Qy      182 LeuSerGlySerLeuValSerLeuHlsCysLeuAlaCysGlyLysSerLeuYerThPro 201
Db      782 CTCTCAGGCTCTCTGCTCTCTGCTGCTGTGTGTGTGGAAAGCTGAGAGACCC 841
Qy      202 ArgValValGlyGlyGlnAlaSerValaspSerTrpProGlnValSerLileGln 221
Db      842 CGTGTGTGTGGGAGAGAGGCTCTGTGTGATCTTGTGCTTGGCAGTCAAGATCTCAG 901
Qy      222 TyrAspLysGlnHlsValCysGlyGlySerLileuapProHlsTrpValLeuThrAla 241
Db      902 TACAGCAAAACAGACGCTGTGTGAGAGGACATCTGTGAGACCCCACTGGGTCTCAGGCA 961
Qy      242 AlaHlsCysPheArgLysHlsThrAspValPheasnrpLysValArgLileGlySerAsp 261
Db      962 GCCCACTGCTTCAAGAAACATACCGATGTGTTCACATGTGAAGGTGGGGAGGCTCAGAC 1021
Qy      262 LysLeuGlySerPheProSerLeuAlaValAlaLysLilellelleGlyPheaspPro 281
Db      1022 AAACGTGGCAGCTTCCCATCTGGCTGTGGCCAGATCATCATTTGAATTCAACCCC 1081
Qy      282 MetTyrProLysaspbnPheHlsAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db      1082 ATGTAACCCAAAGCAATGACATGCGCTCATATAAGCTCAGTCCACTCTTCTCA 1141
Qy      302 GlyThrValArgProLileCysLeuProPhePheaspGlyLuleuThrProAlaThrPro 321
Db      1142 GGCACAGTCAGGCCATCTGTCTGCCCTCTTGAAGAGGACTCACTCAACCCCA 1201
Qy      322 LeuThrLilelleGlyTrpGlyPheThrLysGlnaspGlyLyleMetSeraspLileu 341
Db      1202 CTCTGATCATTTGATGGGCTTTTACGAAGCAATGAGGGGAATGTCTGACACTCTG 1261
Qy      342 LeuGlnAlaSerValGlnValIleaspSerThrArgCysAsnAlaaspAlaTrpGln 361
Db      1262 CTGCAAGGCTCATAGTCAATTCACACACACAGCTGCAATGACAGAGCTTACAG 1321
Qy      362 GlyGlyValThrGlyLysMetMetCysAlaGlyLileProGlyLyleGlyValAspThrCys 381
Db      1322 GGGGAAGTCACCGAAGAATGATGTGTGAGCAATCCCGAAGGGGGTGTGAGACCTGC 1381
Qy      382 GlnGlyAspSerGlyGlyProleuMetLysGlnSerAspGlnTrpHlsValValGlyLile 401

```

```

Db      1382 CAGGTACAGTGTGGGCCCTGATGTACCAATCTGACACAGGATGTGGGCATC 1441
Qy      402 ValSerTrpGlyTyrGlyCysGlyLysProSerThrProGlyValIleThrLysValSer 421
Db      1442 GTTAGCTGGGGCTATGCTCTCGGGGCCCGAGCACTCCGAGGTATACACCAAGTCTCA 1501
Qy      422 AlaTyrLeuaspbnTrpLileTyrAsnValIleTrpLysAlaGlyLuleu 435
Db      1502 GCTTATCTCAACTGATCTACATATGTCTGAAAGGCTGAGCTG 1543

RESULT 13
ID      ADW04403
XX      ADW04403 standard, cDNA, 2104 BP.
AC      ADW04403;
XX      07-APR-2005 (first entry)
XX      DE Human Ovarian specific nucleic acid, Ovr115.
XX      KM Ovarian specific nucleic acid; OSNA; ss; gene; chromosome-11; cytostatic;
XX      KM neoplasm; ovary tumor; Ovr115; SNP; single nucleotide polymorphism.
XX      OS Homo sapiens.
XX      FH Key
XX      FT CDS
XX      FT 233..1546
XX      FT /tag= a
XX      FT /product= "Ovr115"
XX      FT /replace(762,A)
XX      FT variation
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /tag= c
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /replace(824,G)
XX      FT variation
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /tag= d
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /replace(1447,T)
XX      FT /tag= e
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /replace(1572,C)
XX      FT /tag= f
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      FT /replace(1702,A)
XX      FT /tag= g
XX      FT /standard_name= "Single nucleotide polymorphism"
XX      PN W02005005647-A2.
XX      PD 20-JAN-2005.
XX      PP 30-JUN-2004; 2004MO-US021227.
XX      PR 30-JUN-2003; 2003US-0484440P.
XX      PR 30-JUN-2003; 2003US-0484500P.
XX      PA (DIAD-) DIADEXUS INC.
XX      PT Vartanian SF, Macina RA;
XX      DR WPI; 2005-092090/10.
XX      DR P-PSDB; ADW04409.
XX      PT New ovarian specific nucleic acid, useful for detecting, diagnosing,
XX      PT monitoring, staging, imaging and treating ovarian cancer.
XX      PS Example 1; SEQ ID NO 1; 243bp; English.
XX      CC The invention relates to an isolated nucleic acid molecule (ovarian
XX      CC specific nucleic acid). Also included are determining the presence of a
XX      CC ovarian specific nucleic acid (OSNA) in a sample, a vector comprising
XX      CC OSNA, a host cell comprising the vector, producing a polypeptide encoded

```

```
Db      224 GATCTACAGTATCAACCTCTGAACGCTCGATGTCAAACCCCTGGCGAAACCCCGT 283
Qy      22  ILeProMetGluThrPheArgLysValGlyLeProIleIleIleIleAlaLeuLeuSerLeu 41
Db      284 ATCCCATGGAGACTTCAGAAAGGTGGGGATCCCATCATCATAGCACTAGCTAGGCTG 343
Qy      42  AlaserIleIleIleValValLeuIleLysValIleLeuAspLysTrpTrpPheLeu 61
Db      344 GCGAGTATCATATTGGTGTGTCTCTCATCAAGGTGATTTCTGGTAAATATCTACTTCTTC 403
Qy      62  Cys61LysIleProLeuLysPheIleProArgLysGluLeuCysAspArgLysLysLeuAspCys 81
Db      404 TGCAGGAGCCTCTTCACCTCATCCAGAGAGAGAGCTGTGTGACGGAGAGCTGAGCTG 463
Qy      82  ProLeuGluGluAspArgLysLysIleCysValLysSerPheProGluGlyProAlaValAla 101
Db      464 CCTTGGGGGAGAGAGAGAGAGACTGTGTCAAGAGCTTCCCGAAGGGCTGGAGTGGCA 523
Qy      102 ValArgLeuSerLysAspArgSerThrLeuGluValLeuAspSerAlaThrGlyAsnTrp 121
Db      524 GTCCGCTCTTCCAGAGAGCCATCCACTGACAGGTGGTGGACTGGCCACAGGGAACTGG 583
Qy      122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuLysGluThrAlaCysArgGluMet 141
Db      584 TTCTGTCTGTTCGACAACTTCACAAAGCTTCGCTGAGAGAGCTGTAGGACAGATG 643
Qy      142 GlyTrpSerSerLysProThrPheArgAlaValGluIleGlyProAspArgLysLeuAsp 161
Db      644 GGCTACAGACAAACCCACTTTCAGAGCTGTGGAGATTGGCCACAGACAGAGACTGGAT 703
Qy      162 ValValGluIleThrGluAsnSerGlnLysLeuArgMetArgAsnSerSerGlyProCys 181
Db      704 GTTGTGAAATCAGAAACAGCCAGAGAGCTTGGCAAGCGGAAGTCAAGTGGGCTCGT 763
Qy      182 LeuSerGlySerLeuValSerLeuLysCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db      764 CTCTCAGGCTCCCTGTCTCCCTGCACTGTCTTGGAGAGAGCTGAGAGAGCTCAAGCCCC 823
Qy      202 ArgValValGlyGlyGluAlaSerValAspSerTrpProTrpGluValSerIleGln 221
Db      824 CGTGTGTGGGTGGGAGAGAGCTCTGTGATTTGTGGCTTGGCAGGTACAGCTCAG 883
Qy      222 TyrAspLysGlnIleValCysGlyLysSerIleLeuAspProHisTrpValLeuThrAla 241
Db      884 TACGACAAACAGCAGCTCTGTGGAGGAGCATCTCGAGACCCCACTGGGTCTCTCAGGCA 943
Qy      242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db      944 GCCCATCTCTCAGAAACATACGATGTTCCTCACTGGAAGGTGCGGGAGGCTCAGAC 1003
Qy      262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db      1004 AAACCTGGCAGCTTCCATCTCTGGCTGTGGCCAGATCATCATATGAATTCAACCCC 1063
Qy      282 MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db      1064 ATGTACCCCAAGAAATGACATCGCCCTCATGAAGCTGCACTTCCACTCTTCTCA 1123
Qy      302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db      1124 GGCAACAGTCAGGCCCATCTGTCTGCCCTTCTGTATGAGAGACTCACTCCAGCCACCCA 1183
Qy      322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysLysMetSerAspIleLeu 341
Db      1184 CTCTGATCATATGATGGGGCTTTTACGAAGCAGAAATGGAGGAAGATGTCTACATACG 1243
Qy      342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
Db      1244 CTGCGAGGCTCAGTCCAGGCTCATTTGACACACACGGGTGCAATGCAAGCATCGTACCG 1303
Qy      362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyValAlaAspThrCys 381
```

```
Db      1304 GGGGAAGTCACCGAGAAGATGATGTGACAGCATCCCGAAGGGGTGTGACACCTGC 1363
Qy      382 GlnGlyAspSerGlyGlyProLeuMetTrpGlnSerAspGlnTrpHisValIleGlyIle 401
Db      1364 CAGGTGTACAGTGTGTGGCCCTGATGTACCAATTCGACAGTGGCATGTGTGGGATC 1423
Qy      402 ValSerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValTrpThrLysValSer 421
Db      1424 GTTAGCTGGGGCTATGGCTCGGGGGCCGAGACCCCAAGAGTATTCACCAAGTCTCA 1483
Qy      422 AlaTrpLeuAsnTrpIleTrpAsnValTrpLysAlaGluLeu 435
Db      1484 GCCATCTCAACTGAGATCTACATGTCTGGAAGGCTGAGCTG 1525

RESULT 12
ADN04863
ID     ADN04863 standard; cDNA; 2104 BP.
AC     ADN04863;
XX     01-JUL-2004 (first entry)
DT     Antipsoriatic cDNA sequence #646.
DE     ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX     Homo sapiens.
OS     WO2004028479-A2.
PD     08-APR-2004.
XX     25-SEP-2003; 2003WO-US030907.
PF     25-SEP-2002; 2002US-0414006P.
PR     (GERTH ) GENENTECH INC.
PA     Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI,
PI     Wu TD;
XX     WPI; 2004-305105/28.
DR     P-PSDB; ADN04864.
XX     New PRO nucleic acid or polypeptide, useful for preparing a
PT     pharmaceutical composition for diagnosing or treating psoriasis in a
PT     mammal.
XX     Claim 1; SEQ ID NO 1257; 3069bp; English.
XX     The invention relates to novel polynucleotide and polypeptides for
CC     treating psoriasis or a sequence having at least 80% identity to the
CC     above sequences. The nucleic acid is useful for preparing a composition
CC     for diagnosing or treating psoriasis in a mammal. This sequence
CC     corresponds to one of the polynucleotides of the invention.
XX     Sequence 2104 BP; 499 A; 604 C; 577 G; 424 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 6,15e-207 Length: 2104
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 12 Gaps: 0

US-10-803-530-2 (1-435) x ADN04863 (1-2104)
Qy      2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
Db      242 GATCTGACAGTATCAACCTCTGAACGCTCGATGTCAAACCCCTGGCGAAACCCCGT 301
Qy      22  ILeProMetGluThrPheArgLysValGlyLeProIleIleIleIleAlaLeuLeuSerLeu 41
```

Db 310 GTCCGCTCTCCAGGACCGATCCACACTGCTGAGTCTGGCTCGGCGACAGGAACCTG 369
 QY 122 PheSerAlaCysPheAspSerPheThrGluAlaIleuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTGCGCTGTTTCCACAATTCCAGAGAGCTCTGCTGAGACAGCCGTGAGGACGATG 429
 QY 142 GlyTyrSerSerIysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuasp 161
 Db 430 GGCTACAGCAGCAAAACCCATTTCAGAGCTGTGGAGATTGGCCAGACAGAGATCTGGAT 489
 QY 162 ValAlaGluIleThrGluAsnSerGlnIleuAlaGluMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGTAATACAGAAACAGCAGGAGCTTCCATGCGGAATCCAAAGTGGGCGCTGT 549
 QY 182 IeuSerGlySerIeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIleuThrPro 201
 Db 550 CTCTCAGGCTCCCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
 QY 202 ArgValAlaGlyGlyGluGluAlaSerValAspSerTrpProTyrGlnValIleSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGGAGGAGGCTCTGTGGAATCTTGCTTGGCGAGTCCAGCATCAG 669
 QY 222 TyrAspIysGlnHisValCysGlyIysSerIleuAspProHisTrpValIleuThrAla 241
 Db 670 TACGACAAACAGCAGTCTGTGAGAGGAGCATCTGAGACCCCACTGGGTCTTCACGCGCA 729
 QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValAlaGlyIysSerAsp 261
 Db 730 GCCACCTGCTTCAGGAAACATACCGATGTGTTCACCTGAGAGGCGGAGGCTCAGCAGC 789
 QY 262 LysLeuGlySerPheProSerIeuAlaValAlaIysIleIleIleIleGluIleuPheAsnPro 281
 Db 790 AAACCTGGCAGCTCCCATCCCTGCTGTGGCCAAAGATCATCATTAATTCACACCCC 849
 QY 282 MetTyrProIysAspAsnAspIleAlaIleuMetIysLeuGlnPheProIeuThrPheSer 301
 Db 850 ATGTACCCCAAGCAATGACATGCGCCATGACATGCAATGCCATCTTCACATCTTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPheAspGlnIleuThrProAlaThrPro 321
 Db 910 GGACACTCAGGCCCATCTCTGCTCTTCTTGTAGAGAGCTCATCTCAGCAGCACCCCA 969
 QY 322 LeuTrpIleIleGlyTyrGlyPheThrIysGlnAsnGlyGlyIysMetSerAspIleu 341
 Db 970 CTCTGATCATGTGATGGGCTTTACGAGCAATGAGGAGAGATGTCTGACATACCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTCGAGGCTCAGTCCAGGTCATTTGACAGCACCGGTGCAATGACAGATCGTACCG 1089
 QY 362 GlyIleValThrGlyIysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 Db 1090 GGGGAAGTCCACGAGAGATGATGTGTGAGGCGATCCCGAAGGGGCTGTGACACCTGCG 1149
 QY 382 GlnGlyAspSerGlyIysProIeuMetIysGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1150 CAGGGGACAGTGTGGGCGCCCGATGTCATTCACATGACATGGCAGTGGTGGGCTC 1209
 QY 402 ValSerTrpGlyTyrGlyIysCysGlyIysProSerThrProGlyValTyrThrIysValSer 421
 Db 1210 GTTAGCTGGGCTATGCTGCGGGGCGGAGCACCCAGAGATTTACACCAAGGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIysAlaGluIleu 435
 Db 1270 GCTTATCTCACTGATCTACATGTCTGGAAGGCTGAGCTG 1311
 RESULT 11
 AAI72976
 ID AAI72976 standard; cdna; 2081 BP.
 XX
 AC AAI72976;
 XX
 DT 21-AUG-2002 (first entry)

XX CXA8 preferred cDNA.
 DE Gene; colorectal cancer; CGA7; CJA8; modulating protein; screening;
 XX drug candidate; vaccine; ss.
 KW Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 215..1528
 FT /tag= a
 XX /product= "CJA8"
 PN US2002042067-A1.
 XX
 PD 11-APR-2002.
 XX
 PF 08-MAY-2001; 2001US-00851588.
 XX
 PR 17-AUG-2000; 2000US-00642252.
 PR 06-SEP-2000; 2000US-00656002.
 XX
 PA (MACK/) MACK D.
 PA (GISH/) GISH K.C.
 PA (WILS/) WILSON K E.
 XX
 PI Mack D, Gish KC, Wilson KE;
 DR WPI: 2002-453647/48.
 DR P-PSDB: AAG79359.
 XX
 PT Screening drug candidates for treating colorectal cancer, comprises
 PT determining the effect of the candidate on the expression profile gene of
 PT CGA7 or CJA8.
 XX
 PS Claim 69; Fig 7; 40pp; English.
 XX
 CC The sequences given in AAI72973-76 encode the colorectal cancer proteins,
 CC CGA7 and CJA8. The CGA7 and CJA8 proteins are colorectal cancer
 CC modulating proteins and have been mapped to chromosomes 2 (CGA7) and 11
 CC (CJA8). These sequences may be used in the method of the invention for
 CC screening drug candidates. The method comprises adding a drug candidate
 CC to a cell that expresses an expression profile gene encoding CGA7, CJA8
 CC or fragments and determining the effect of the drug candidate on the
 CC expression of the expression profile gene. The new methods are used to
 CC screen bioactive agents for the ability to bind to or modulate the
 CC activity of CGA7 or CJA8 and evaluate the effect of a candidate
 CC colorectal cancer drug. An antibody to CGA7 or CJA8 can inhibit the
 CC activity of CGA7 or CJA8, respectively, and is used to screen for an
 CC agent that can interfere with the binding of CGA7 or CJA8 to the
 CC antibody. The antibody can be used to treat colorectal cancer. The
 CC colorectal cancer tissue, where the therapeutic group is a cytotoxic
 CC agent or a radiolabeled. Antisense molecules are used to inhibit
 CC colorectal cancer in a cell. Nucleic acid segments encoding CGA7 or CJA8
 CC are used in a biochip. CGA7, CJA8 or a nucleic acid encoding it are used
 CC to elicit an immune response. CGA7 or CJA8 is used to determine the
 CC prognosis of an individual with colorectal cancer. Nucleic acid encoding
 CC CGA7 or CJA8 can be used in vaccines
 XX
 SQ Sequence 2081 BP; 484 A; 597 C; 576 G; 424 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6,06e-207 Length: 2081
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 6 Gaps: 0
 US-10-803-530-2 (1-435) x AAI72976 (1-2081)
 QY 2 AspProAspSerAspGlnProIeuAsnSerIeuAspValIysProIeuAlaGlyIysProArg 21

QY 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValIleuThrAla 241
 DB 670 TACGACAAACGACGCTGTGTGGAGGAGCATCCGAGCCCCACTGGGTCTCTACGGCA 729
 QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
 DB 730 GCCCATGCTCTCAGGAACATACGATGTGTCAACTGGAGAGTGGCGGACGGCTCAGAC 789
 QY 262 LysIleuGlySerPheProSerIleuAlaValIleIleIleGluPheAsnPro 281
 DB 790 AAACGGGACACTTCCCATCCCTGCTGTGGCCAAAGATCATCATTAATTAACACCC 849
 QY 282 MetTrpProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 850 ATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCAGTCCCATCTTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluIleuThrProAlaThrPro 321
 DB 910 GGCAACAGTCAGGCCCATCTGTCTGCCCTTTGTATGAGAGCTCATCTCCAGCCACCCCA 969
 QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleu 341
 DB 970 CTCTGGATCATTTGGATGGGCTTTACGAAGCAGATGAGGGAAGATGCTGACACTACTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
 DB 1030 CTGCAAGGCTTCACTTCAGCTCATTTGACGACACCGCTGCATGACAGACAGTCCGTAACG 1089
 QY 362 GlyLysValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 DB 1090 GGGGAAGTCACCGAAGATATGTGTCCAGGCATCCCGGAAGGGGGTGTGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyLysProLeuMetCysGlnSerAspGlnTrpHisValAlaGlyIle 401
 DB 1150 CAGGCTGACAGTGTGGGCCCCCTGATGTACCAATCTGACCACTGGCATGTGGTGGGCATC 1209
 QY 402 ValSerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValIleThrLysValSer 421
 DB 1210 GTTACCTGGGCTATAGCTGTGGGGGGCCCGAGACCCCAAGGATATACCAAGGCTCTCA 1269
 QY 422 AlaTrpLeuAsnTrpIleTrpAsnValTrpLysAlaGluLeu 435
 DB 1270 GCCTATCTCACTGATCTACATGTCTGGAAGCTGAGCTG 1311
 RESULT 10
 ADU25739 ID ADU25739 standard; cDNA; 1314 BP.
 XX AC ADU25739;
 XX AC
 DT 27-JAN-2005 (first entry)
 XX AC
 DE CDNA encoding human TMPRSS4 protease.
 XX
 KW Human: transmembrane serine protease 4; TMPRSS4: cardiovascular disorder;
 KW endocrinological disease; gastroenterological disease;
 KW urological disorder; metabolic disease; cancer; respiratory disease;
 KW dermatological disease; cardiact; antidiabetic; antidiabetic;
 KW hyperosmotic; antidiabetic; laxative; nephrotoxic; nootropic;
 KW osteopathic; neuroprotective; antiparkinsonian; cytoskeletal; gene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1314
 FT /tag= a
 FT /product= "TMPRSS4 protease"
 XX
 XX WO2004097034-A2.
 XX
 PD 11-NOV-2004.
 XX

PF 20-APR-2004; 2004MO-EP004164.
 XX
 PR 02-MAY-2003; 2003EP-00010004.
 XX
 PA (FARB) BAYER HEALTHCARE AG.
 XX
 PI Golz S, Brueggemeier U, Geerts A, Polej S;
 DR WPI: 2004-804771/79.
 DR P-PSDB: ADU25740.
 XX
 PT Screening for therapeutic agents, useful for treating cancer,
 PT cardiovascular, endocrinological, urological, metabolic, respiratory, or
 PT dermatological diseases, comprises contacting a test compound with a
 PT transmembrane serine protease 4.
 XX
 PS Disclosure; SEQ ID NO 1; 134bp; English.
 XX
 SS The invention relates to a method of screening for therapeutic agents
 CC useful in the treatment of human disease. The method comprises contacting
 CC a test compound with a transmembrane serine protease 4 (TMPRSS4)
 CC polypeptide or polynucleotide, or determining the activity of a TMPRSS4
 CC polypeptide at a certain concentration of a test compound or in the
 CC absence of the test compound. The method is useful for diagnosing,
 CC treating, and preventing cardiovascular disorders (e.g. myocardial
 CC infarction, arrhythmias, hypertensive or peripheral vascular diseases, or
 CC atherosclerosis), endocrinological diseases (e.g. diabetes mellitus,
 CC Zollinger-Ellison syndrome, Addison's disease, Cushing's syndrome,
 CC hyperaldosteronism, or thyrotoxicosis), gastroenterological diseases (e.g.
 CC gastritis, chronic inflammatory disease, Crohn's disease, diarrhoea, or
 CC constipation), urological disorders (e.g. renal disease,
 CC glomerulopathies, nephritis, or obstructive uropathies), metabolic
 CC diseases (e.g. obesity, Lesch-Nyhan syndrome, osteoporosis, Alzheimer's
 CC disease, Parkinson's disease, or Fanconi's syndrome), cancer, respiratory
 CC diseases (e.g. asthma or chronic obstructive pulmonary disease) and
 CC dermatological diseases (e.g. psoriasis, ichthyosis, keratosis pilaris,
 CC psoriasis, pityriasis rosea, or dermatitis). This sequence encodes human
 CC TMPRSS4 protease.
 XX
 SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 3,3e-207 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 13 Gaps: 0
 US-10-803-530-2 (1-435) x ADU25739 (1-1314)
 QY 2 AspProAspSerArgGlnProLeuAsnSerIleuAspValLysProLeuArgLysProArg 21
 DB 10 GATCTGACAGTGTATCAACCTCTCAACAGCTTCATGCAACCCCTGGCAACCCCGT 69
 QY 22 IleProMetGluTrpPheArgLysValGlyIleProIleIleIleAlaLeuSerIleu 41
 DB 70 ATCCCATGAGACCTTACGAAAGTGGGATCCCATCATCATATGACATCTAGGCTG 129
 QY 42 AlaSerIleIleIleValAlaValIleuIleValIleuAspLysTrpPheLeu 61
 DB 130 GCGAGTATCATCTGTGGTGTCTCTATCAAGATGATTCGATTAATATCTACTCTC 189
 QY 62 CysGlyLysProLeuHisPheIleProArgLysGlnLeuCysAspGlyLysLeuAspCys 81
 DB 190 TGGCGGAGCCTCTTCACTTATCCCGAGGAAGAGCTGTGTGACGAGACTGACTGT 249
 QY 82 ProLeuGlyLysAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
 DB 250 CCTTTGGGAGGAGACGAGACACTGTGTCAAGACTTCCCGAAGGCTGTGAGTGA 309
 QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121

Db	1090	GGGGAAGTCCAGGAAGATGATGTGCAGGACATCCCGAAGGGGTGTGACACATGTC	1149
Qy	382	GlnAlYApSerSrgIyGIYProLeuMeTYrGInSerAspGILThrHISValValGlyLe	401
Db	1150	CAGGTGACAGTGTGGGCCCTGATGTACCATCTGACCAAGTGGCATGTGTGGGCATC	1209
Qy	402	ValSerTrpGILYrYrGIYCYrGIYGIYProSerThrProGILYValTYrThrLYrValSer	421
Db	1210	GTYAGCTGGGGCTATAGCTCGGGGGGCCGACGCCAGCATGATATACCAAGGTCTCA	1269
Qy	422	AlaTYrLeuAsnTrpILETYrAsnValLTTrpValAlaGluLeu	435
Db	1270	GCCATCTCAACTGCATCTTCAATGTCATGGAGGCTGAGCTG	1311
RESULT 9			
ADN39460	ID	ADN39460 standard; cDNA, 1314 BP.	
XX	AC	ADN39460;	
XX	DT	17-JUN-2004 (first entry)	
XX	DE	Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:A60.	
XX	KW	Human; differential expression; cancer; angiogenic disorder;	
KW	KW	fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;	
KW	KW	inflammatory disease; autoimmune disease;	
KW	KW	retinal neovascularisation syndrome; scarring; uterine fibroid;	
KW	KW	detection; diagnosis; prognosis; drug screening; drug targeting;	
KW	KW	wound healing; contraception; cytostatic; cardant; immunomodulatory;	
XX	OS	vulnerable; gene therapy; vaccine; gene; ss.	
XX	XX	Homo sapiens.	
PN	XX	WO2003042661-A2.	
XX	PD	22-MAY-2003.	
XX	PF	13-NOV-2002; 2002WO-US036810.	
PR	13-NOV-2001; 2001US-0350666P.		
PR	21-NOV-2001; 2001US-0332464P.		
PR	29-NOV-2001; 2001US-0334393P.		
PR	03-DEC-2001; 2001US-0335394P.		
PR	14-DEC-2001; 2001US-0340376P.		
PR	08-JAN-2002; 2002US-0347211P.		
PR	10-JAN-2002; 2002US-0347349P.		
PR	08-FEB-2002; 2002US-0355250P.		
PR	13-FEB-2002; 2002US-0356714P.		
PR	20-FEB-2002; 2002US-0359077P.		
PR	09-MAR-2002; 2002US-0368809P.		
PR	04-APR-2002; 2002US-0370110P.		
PR	12-APR-2002; 2002US-0372246P.		
PR	05-JUN-2002; 2002US-0386614P.		
PR	16-JUL-2002; 2002US-0396839P.		
PR	22-JUL-2002; 2002US-0397775P.		
PR	22-JUL-2002; 2002US-0397845P.		
PR	09-SEP-2002; 2002US-0409450P.		
XX	PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX	PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Heveri PA,	
PI	MacK DH, Murray R, Watson SR, Wilson KE, Zlotnik A,		
XX	DR	WPI; 2003-468649/44.	
DR	P-PSDB; ADN39461.		
XX	PT	Determining the presence or absence of a pathological cell in a patient,	
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting		
PT	a nucleic acid in a biological sample.		
XX	PS	Claim 8; SEQ ID NO A60; 1385pp; English.	

XX Homo sapiens.
OS
XX NO2003042661-A2.
PN
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US0366810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-035394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 06-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
XX WPI; 2003-468649/44.
DR P-PSDB; ADN38816.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX
XX Claim 8; SEQ ID NO 133; 1385bp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
XX Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
SO
Alignment Scores:
Pred. No.: 3.3e-207 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 11 Gaps: 0
US-10-803-530-2 (1-435) x ADN38815 (1-1314)
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValValProLeuArgLysProArg 21

Db 10 GATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGCAAAACCCCTGCGCAACCCCGT 69
Qy 22 ILeProMetGlnThrPheArgLysValGlyIleProIleIleIleIleIleLeuSerLeu 41
Db 70 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATCATGACATGAGCCTG 129
Qy 42 ALaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIleVal 61
Db 130 GCGAGTATCATCATTTGGTGTCTTCATCATCATCATCATCATCATCATCATCATCATCATCATCTTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyIleLeuAspCys 81
Db 190 TGGCGGAGCCTCTCCATCTTCATCCAGAGAACAGCTGTGTGACGGAGAGCTGACTGT 249
Qy 82 ProLeuGlyIleAspGlyIleHisCysValLysSerPheProGlyIleProIleValIle 101
Db 250 CCTTTGGGGGAGAGCAGAGGACCTGTGTCAAGAGCTTCCCGAAGGGCCTGACAGTGGCA 309
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTTCAGAGACCATTCACACTGAGAGTCTGAGCTGGCCACAGGAACTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet 141
Db 370 TTCTCGCTGTGTTTCGAACTTCACAGAGCTCTCGTGAAGACAGCTGTAGGACAGATG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGACGAAACCCACTTCAGAGCTGTGAGATGGCCAGACCCAGAGACTGTGAT 489
Qy 162 ValValGlnIleThrGlnAsnSerGlnGlnLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTGTGTAATCAGAAAGAACAGCAGAGAGCTGTGACAGCCGAACCTCAAGTGGGCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCACAGCTCTCCGCTGCTCTCCCTGACATGCTTCTCGTGGAGAGAGCTGAAGACCCCC 609
Qy 202 ArgValValGlyIleGlnIleValSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGGGTGGTGGGAGAGGCTCTGTGATTTCTGGCTTGGCAGGTGAGATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACACAAACGACAGCTGTGGAGGAGCATCTGAGCCCACTGGGCTCTCAAGGCA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCACCTGCTTCAGAAACATACCGATGTTCACATGGAAGTGGGAGGCTCAGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGlnPheAsnPro 281
Db 790 AAATCGGGCAGCTTCCATCCCTGCTGTGGCCAAAGATCATCATCATCATCATCATCATCATCAT 849
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTAACCCCAAGAACATGATGCTCTAGAGCTGCAAGTCCCACTCACTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGlnIleLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAAGCCCATGTGCTGCCCTCTTTATATGAGAGCTCACTCCAGCCACCCCA 969
Qy 322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
Db 970 CTCTGATCATTTGATGGGCTTTTACGAGCAGATGAGAGGAAATGTCTGACATCTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGCAAGGCTGATGCTCAGAGCTATTCAGACACACGCTGCATTCAGAGAGTGCATCAG 1089
Qy 362 GlyIleValIleThrGlyLysMetMetCysAlaGlyIleProGlyIleGlyValAspThrCys 381

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
XX WPI; 2003-468649/44.
DR P-PDSB; ADN39473.
XX

PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX
XX
PS Claim 8; SEQ ID NO A72; 1385bp; English.

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularization syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,3e-207 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
Gaps: 0

US-10-803-530-2 (1-435) x ADN39472 (1-1314)

QY 2 AspProAspSerArgGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
Db 10 GATCTCGACAGTATCATCACTTGAACAGCGCTCGATGTCAACCCCTGGGCAAAACCCCT 69
QY 22 IleProMetGlnThrPheArgIysValGlyIleProIleIleIleAlaLeuSerIleu 41
Db 70 ATCCCATGAGACCTTCAAGAAAGTGGGAGTCCCATCATCATACACTGAGAGCTG 129
QY 42 AlaSerIleIleIleValValIleuIleValIleValIleAspIysPheIleu 61
Db 130 GCGAGTATCATCTGTGGTGTCTCCATCATCAAGTATCTGATTAATTAATTAATCTCC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspIyGluLeuAspCys 81
Db 190 TGGGGGACCCCTCTCCACTTCATCCCGAAGAGACACTGTGAGAGAGAGCTGACTGT 249
QY 82 ProLeuGlyGluAspGluGlnHisCysValIysSerPheProGluGlyProAlaValAla 101
Db 250 CCTTGGGGAG 309
QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTTTGAACAACCTTCAAGAGAGCTCTGCTGAGAGAGAGAGAGAGAGAG 429
QY 142 GlyTyrSerSerIysProThrPheArgIysValGlyIleGlyProAspGlnAspLeuAsp 161
Db 430 GCGTACAG 489

QY 162 ValValGluIleThrGluAsnSerGlnIleuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATTCACAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 549
QY 182 LeuSerGlySerIleuValSerLeuHisCysLeuAlaCysGlyIysSerLeuThrPro 201
Db 550 CTCTCAGGCTCCCTGTGTCTCCCTGACATGTCTTGGCTGTGGAGAGAGAGAGAGAGAG 609
QY 202 ArgValValGlyGluGlnGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGGTGGGAG 669
QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGACAAACACACAGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTCAGAGAAACATACCATGTGTTCACCTGAAAGTGGCGGAGAGAGAGAG 789
QY 262 LysLeuGlySerPheProSerIleuAlaValAlaValIleIleIleGluPheAsnPro 281
Db 790 AAACCTGGGAG 849
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAAATGACATGCGCTCATGAAAGCTGAGAGAGAGAGAGAGAGAG 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGlnIleuThrProAlaThrPro 321
Db 910 GGCACAGTCAG 969
QY 322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyIysMetSerAspIleu 341
Db 970 CTCTGAGATCATGATGAG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGACAGGCGTACGTCAGAGTATTCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1089
QY 362 GlyGluValThrGlyLysMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db 1090 GGGAGAGTCAACCGAAG 1149
QY 382 GlnIysAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
Db 1150 CAGGATGACAGTGTGGGAG 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIysProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGGCTGCGGGGCGGAGACCCAGAGAGATATACACAGAGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTTATCTCAACTGAGATCAATATGTGGAAGGCTGAGAGCTG 1311

RESULT 8
ADN38815
ID ADN38815 standard; cDNA; 1314 BP.
XX
AC ADN38815;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:133.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.

Alignment Scores:

Pred. No.: 3,36-207 Length: 1314
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 11 Gaps: 0

US-10-803-530-2 (1-435) x ADN39661 (1-1314)

QY 2 AspProaspSeraspGlnProleuasnserleuaspValylsProleuarglyseProarg 21
 Db 10 GATCTGACAGTGTATCAACCTCTGAACAGCCTTCATGCAAAACCCCTCGCAAAACCCCTG 69
 QY 22 IleProMetGluThrPheArgLyValGlyIleProIleIleIleAlaLeuLeuSerleu 41
 Db 70 ATCCCCATGAGACCTTCAGAAAGGTGGGAGATCCCATCATCATATGACACTAGAGCTG 129
 QY 42 AlaSerIleIleIleValValValLeuIleValIleLeuAspIleValIleValIleVal 61
 Db 130 GCGAGTATCATATTGGTGTGCTCTCATCAAGGTGATTCGATTAATTAATTAATTAATTC 189
 QY 62 CysGlyGlnProleuAsnHisPheIleProarglyseGlnleuGlyAspGlyGlyLeuAspCys 81
 Db 190 TCGCGGAGCCTCTCTCACTTCACTCCGAGAAAGCAGCTGTGTGACGAGAGAGCTGACTGT 249
 QY 82 ProleuGlyGlyAspGlyGlnHisPheValylsSerPheProGlyGlyProAlaValAla 101
 Db 250 CCTTTGGGGGAGAGAGAGAGACATGTGTCAAGAGCTTCCCGAAAGGGCCTGACGTGCA 309
 QY 102 ValArgLeuSerLyAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
 Db 310 GTCCGCCCTCTCCAGAGACCGATCCACACTGCAAGGTGTGAGCTCGGCCACAGGGAACTGG 369
 QY 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTGTGCTGTGTGACAACTTCAAGAAAGCTGTGCTGAGAGACGCTGTAGAGAGATG 429
 QY 142 GlyTyrSerSerLyPProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACACACGAAACCACTTTCAGACCTGTGAGATTTGGGCCGACACAGATGTGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 Db 490 GTTGTGAATTCACAGAAACAGCCAGAGAGCTTCCAGATGCGGAATCAAGTGGGCGCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLySerLeuIleThrPro 201
 Db 550 CTCTAGGCTCTCTGTGTCTCTCTGCACTGTCTGTGTGGGAAAGAGCTTGAAAGACCCCT 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
 Db 610 CGTGTGTGTGTGGGAGAGAGGCTCTGTGATTTCTTGCCCTTGGCAGGTCCGATCCAG 669
 QY 222 TyrAspLyGlnHisValCysGlyGlySerIleLeuAspProHisIleTrpValLeuThrAla 241
 Db 670 TACGACAAACGACAGTCTGTGAGAGGAGCATCCGGAACCCCATGTGGTCTCTCAAGGGA 729
 QY 242 AlaHisCysPheArgLyHisIleThrAspValPheAsnTrpLyValArgAlaGlySerAsp 261
 Db 730 GCCCACTGCTTCAGAAACATACCGATGTGTTCACTGGAAGGTGCGGGCAGGCTCAAGC 789
 QY 262 LysLeuGlySerPheProSerIleuAlaValAlaValIleIleIleGluPheAsnPro 281
 Db 790 AAACGTGGGAGCTTCCCATCTCTGTGCTGGCCAAAGATCATATCATTAATTAATTAATTA 849
 QY 282 MetTyrProLyAspAsnAspIleAlaLeuMetLyLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTACCCCAAGACATGACATGACGCGCCCATGAAGGTGCAATTCACATCACTTCTTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlyGluLeuThrProAlaThrPro 321
 Db 910 GGACACATCAGGCCCATGTGTGCTGCTTCTTTGATGAGAGACTCATCTCAACCAACCCCA 969

QY 322 LeuTrpIleIleIleGlyTrpGlyPheThrLyseGlnAsnGlyLyLysMetSerAspIleLeu 341
 Db 970 CTCTGATCATTTGGATGGGCTTTTACAGACAGATGAGAGGAAAGATCTCGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTCGACGCGTCAGTCCAGGTCAATGACAGCACAGGTGCAATGACAGATGCGTACAG 1089
 QY 362 GlyGluValThrGlyLysMetMetCysAlaGlyIleProGlyGlyGlyValAspThrCys 381
 Db 1090 GGGGAAATCCAGAGAGATGATGTGCAAGGCTATCCCGGAAGGGGTGTGGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyLyProleuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
 Db 1150 CAGGTACAGTGTGGGCCCTGTGATTAACCAATCAACAGTGCATGTGTGGGCAATC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLyValSer 421
 Db 1210 GTTAGCTGGGCTATGGCTGCGGGGCCCGAGCACCCCAAGATATACCAAGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLyAlaGluLeu 435
 Db 1270 GCTATCTCAACTGATCTCAATGTCTGGAAGGCTGAGCTG 1311

RESULT 7
 ADN39472
 ID ADN39472 standard; cDNA; 1314 BP.
 AC ADN39472;
 DT 17-JUN-2004 (first entry)
 DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:472.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytosolic; cardiant; immunomodulatory;
 KW vunerary; gene therapy; vaccine; gene; ss.
 OS Homo sapiens.
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0355250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368099P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

Db 310 GTCCGCTTCCAGGACCGATCCACAGCAGTCTGAGCTCCGACAGGAACTCG 369
 QY 122 PheSerAlaCySPheAspAspPheThrGluAlaGluThrAlaCysArgGlnMet 141
 Db 370 TTCTGTGCTGTTTGGACAACTTCACAGAGCTCTCGCTGAGACAGCTGTAGGCAATG 429
 QY 142 GilyTySerSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 Db 430 GGCTACAGCAGCAAACTTCAGAGCTGTGAGATTGGCCAGACCAAGATCTGGAT 489
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuAlaGluMetLysAsnSerSerGlyProCys 181
 Db 490 GTTGTGAAATCACAGAAACAGCAGAGCTTCCAGTCAGGAACTCAAGTGAGCCCTGT 549
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
 Db 550 CTCTCAGGCTCCCTGTGCTCTCTGCTGACATGCTCTGTGAGAAAGAGCTGAAAGACCCC 609
 QY 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProTyrGlnValSerIleGln 221
 Db 610 CGTGTGTGGTGGGAGAGAGCCCTGTGTGATTCTTGAGCTTGAGAGTCAAGTATCAG 669
 QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
 Db 670 TACACAAACAGCAGCTCTGTGAGAGGAGCATCTGAGACCCCTGAGTCTTCAAGGCA 729
 QY 242 AlaHisCysPheAspGlyLysHisThrAspValPheAsnTrpLysValAlaGlySerAsp 261
 Db 730 GCCACAGCTTCAAGAAACATACCGATGTGTTCAACTGAGAAAGTGCGGAGGCTCAGAC 789
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
 Db 790 AAATGTGGAGCTTCCATCCCTGCTGTGGCCAGATCATCATTAATTAATCAACCCC 849
 QY 282 MetCysProLysAspAspAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 Db 850 ATGTAACCCCAAGACATGACATCGCCCTCATGAGCTGAGTCCCACTTTCTCTCA 909
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThrPro 321
 Db 910 GGACAGCTGAGGCCATCTCTGCTCTTGTGATGAGAGCTCACTCCAGCACCCCA 969
 QY 322 LeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 Db 970 CTCGTGATCATGTGAGGCTTTACGAGACAGATGAGAGGAGATGCTGACATCTG 1029
 QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 Db 1030 CTGCAGGCGTCATCCAGGTCATTCACACACGCGTGCATGACAGATGCGTACCG 1089
 QY 362 GlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
 Db 1090 GGGGAGTCAACCGAAGATGATGTGTGAGAGGATCCGGAAGGGGTGTGACACCTGC 1149
 QY 382 GlnGlyAspSerGlyGlyProLeuMetLysGlnSerAspGlnTrpHisValValGlyIle 401
 Db 1150 CAGGGTGACAGTGTGGGCCCTGATGTACATCTGACCAITGGCAITGGTGGGCAATC 1209
 QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
 Db 1210 GTTACGTGGGCTATGTGCTGCGGGGCCGAGCACCCAGAGGTATACACCAAGTCTCA 1269
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTyrLysValGluLeu 435
 Db 1270 GCCTATCTCACTGATCTCAATATGTGGAAGGCTGAGCTG 1311
 RESULT 6
 ADN39661
 ID ADN39661 standard; cDNA; 1314 BP.
 XX AC
 XX ADN39661;
 XX

DT 17-JUN-2004 (first entry)
 XX
 DB Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: C33.
 XX
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularization syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW
 XX
 OS Homo sapiens.
 XX
 EN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR 29-NOV-2001; 2001US-0334393P.
 PR 03-DEC-2001; 2001US-0335394P.
 PR 14-DEC-2001; 2001US-0340376P.
 PR 08-JAN-2002; 2002US-0347211P.
 PR 10-JAN-2002; 2002US-0347349P.
 PR 08-FEB-2002; 2002US-0356250P.
 PR 13-FEB-2002; 2002US-0356714P.
 PR 20-FEB-2002; 2002US-0359077P.
 PR 29-MAR-2002; 2002US-0368097P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 16-JUN-2002; 2002US-0386614P.
 PR 05-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 PI Afar D, Aziz N, Ginsburg NM, Gish KC, Glynn R, Hevezi PA,
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
 XX
 DR WPI; 2003-468649/44.
 DR P-PSDB; ADN39878.
 XX
 PT Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 PS Claim 8; SEQ ID NO C33; 1385pp; English.
 XX
 SS The invention relates to nucleic acids and proteins (ADN39683-ADN40064)
 XX whose expression is upregulated or downregulated in specific cancers or
 XX other diseases such as angiogenic or fibrotic disorders, and to methods
 XX of determining the presence or absence of a pathological cell in a
 XX patient by detecting a nucleic acid at least 80% identical to those of
 XX the invention or by detecting a polypeptide of the invention. The
 XX invention also relates to expression vectors and host cells comprising a
 XX nucleic acid of the invention; antibodies which specifically bind a
 XX polypeptide of the invention; use of such antibodies for drug targeting;
 XX and methods of screening for modulators of activity or expression of the
 XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
 XX antibodies and methods are useful for diagnosing, prognosing and treating
 XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
 XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 XX neovascularization syndromes, scarring and uterine fibroids. They may
 XX also be useful in wound healing and in contraception. The present
 XX sequence represents a nucleic acid sequence of the invention.
 SO Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

QY 262 LysLeuGlySerPheProSerLeuAlaValAlaIleIleIleIleGluPheAsnPro 281
DB 790 AAACGGGACAGCTCCATCCCTGGCTGGCGAAGATCATCATATTGAATCAACCCC 849
QY 282 MetTyrProIlyAspAsnAspIleAlaLeuMetIlyLeuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAGCAATGACATCCCTCATGAGCGCAGTTCCTCACTTCTTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluIleuThrProAlaThrPro 321
DB 910 GGCACAGTCAGGCCCATCTGTCTGCTCTTGTATGAGAGCTCATCTCCAGCCCA 969
QY 322 LeuTrpIleIleGlyTyrGlyPheThrIlyGlnAsnGlyIlyMetSerAspIleLeu 341
DB 970 CTCCTGATCATTTGATGGGGCTTTACGAGCAGATGAGGGAAGATGCTGACATCTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaTyrGln 361
DB 1030 CTGCGAGCGCTACAGTCAGTCATGACAGCACCGCTCAATGACAGATGCGTACCG 1089
QY 362 GlyGlyValThrGluValMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
DB 1090 GGGGAGGTCACCGAAGATGATGTGTCCAGCGATCCGGAAGGGGGTGTGACACCTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheValIleGlyIle 401
DB 1150 CAGGCTGACAGTGTGGGGCCCTGTATGATGATGACATGACAGTGTGCTGGGCTC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIlyProSerThrProGlyValIlyThrIlyValSer 421
DB 1210 GTTAGCTGGGCTATAGCTGGCGGGGCCGAGCACCCGAGAGTATACACCAAGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIlyAsnIleGluLeu 435
DB 1270 GCTATCTCACTGATCTCAATGTCTGGAAGGCTGAGCTG 1311
RESULT 5
ADN39512
ID ADN39512 standard; cDNA; 1314 BP.
AC ADN39512;
XX 17-JUN-2004 (first entry)
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:112.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.

PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hovezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR P-PSDB; ADN39513.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO 112; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,3e-207 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 11 Gaps: 0
US-10-803-530-2 (1-435) x ADN39512 (1-1314)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIlyProLeuArgProArg 21
DB 10 GATCCTGACAGTGAATCAACTCTGAACAGCTCGATGCAAACTCCGCAAACTCCGT 69
QY 22 IleProMetGluThrPheArgIlyValGlyIleProIleIleIleAlaLeuSerLeu 41
DB 70 ATCCCATGAGACCTTCAAGAAAGTGGGATCCCAATCATCATGACATCTAGGCTG 129
QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleVal 61
DB 130 GCGAGTATCATCATTTGCTGTCTCTCATCAAGGATTCGATTAATTAATTAATCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIlyGlnLeuCysAspGlyIlyLeuAspCys 81
DB 190 TGGCGGCGAGCTCTCCATCTCATCCCGAAGCAGCTGTGTGACGAGAGCTGACCTGT 249
QY 82 ProLeuGlyIlyAspGluGluHisCysValIlySerPheProGluGlyProAlaValAla 101
DB 250 CCTTTGGGAG 309
QY 102 ValArgLeuSerIlyAspArgSerThrLeuGlnValLeuAspSerValThrGlyAsnTrp 121

Db 970 CTCGTGATCATGATGGGCTTTACAGAGAGAGATGAGATGTCTGACATVACTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAlaIleTyrGln 361
Db 1030 CTGAGGGGTGATGCTCAGGATCATGACAGACACGGGTCATGACAGCATGGCTACAG 1089
QY 362 G1yGlnValThrGlyuMetMetCysAlaGlyIleProGlnuGlyIleValAspThrCys 381
Db 1090 GGGGAAGTCACCGAAGATGATGTGACGAGCATCCCGAAGGGGTGTGACACCTGC 1149
QY 382 GlnG1yAspSerGlyGlyProLeuMetTyrGlnSerAspGlnThrPheValAlaGlyIle 401
Db 1150 CAGGGTACAGTGGTGGCCCTGATGACCAATCTGACAGTGGCATGTGGGCAATC 1209
QY 402 ValSerThrGlyIleTyrGlyCysGlyIleProSerThrProGlyValIleTyrThrValSer 421
Db 1210 GTTAGCTGGGGCTATGCTGCGGGGCCGAGACCCCGAGGATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnThrIleTyrAsnValIlePheValAlaGlnLeu 435
Db 1270 GCCTATCTCAACTGAGATCTACAAATGTCTGAAAGGCTGAGCTG 1311

RESULT 4
ADB80524
ID ADB80524 standard; DNA; 1314 BP.
XX
AC ADB80524;
XX
DT 04-DEC-2003 (first entry)
XX
DE Ovarian cancer-associated transcript #45.
XX
KW cytosolic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.
XX
OS Homo sapiens.

Key Location/Qualifiers
FT CDS 1..1314
FT /tag=a
XX
PN WO2002102235-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (BOSB-) BOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Gish KC;
XX
DR WPI; 2003-167431/16.
DR P-PSDB; ADB80525.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient; comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS Claim 10; Page 305; 332pp; English.
XX
CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or

CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX

SQ Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;
Alignment Scores:
Pred. NO.:
Score: 3.3e-207 Length: 1314
Percent Similarity: 2337.00 Matches: 434
Best Local Similarity: 100.00% Conservative: 0
Query Match: 99.79% Mismatches: 0
DB: Indels: 0
Gaps: 0

US-10-803-530-2 (1-435) x ADB80524 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIleProLeuAlaGlyLysProArg 21
Db 10 GATCTGACAGATGATCACTTGAACAGCTCGATGCAAAACCCCTGGCGAAACCCCGT 69
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCCATGAGAACCTTCAGAAAGTGGGAGATCCCATCATCATAGCACTACAGCTG 129
QY 42 AlaSerIleIleIleValValIleuIleLysValIleLeuAspLysTyrTyrPheLeu 61
Db 130 GCGATATACATCATTTGGTGTGCTCATAGGTGATTTGGATTAATACTCTTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlyLysAspCys 81
Db 190 TCGGGGAGCTCTCCACTTCATCCGAGAGACAGCTGTGTACGAGAGCTGAGCTGT 249
QY 82 ProLeuGlyGluAspGlyGluHisCysValLysSerPheProGlnuGlyProAlaValAla 101
Db 250 CCTTGGGGGAGAGCAGAGAGCACTGTGTCAAGAGCTTCCCGAAGGGCTGAGTGGCA 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleuAspSerLThrGlyAsnTrp 121
Db 310 GTCCGCTCTTCAAGAGCCATTCACAGAGCTGTGACACAGCTGGCCACAGGAACTGG 369
QY 122 PheSerAlaCysPheAspAsnThrGlnAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCCTGCTGTTTCGACAACTTCACAGAGCTGTGACACAGCTGTGAGCATG 429
QY 142 G1yTyrSerSerLysProThrPheAlaGlyAlaGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGTACAGCAGCAAAACCACTTCAAGCTGTGAGATTGGGCCACAGCATCTGGAT 489
QY 162 ValValGlnuIleThrGlnuSerGlnuLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATATCAGAAACAGCAGCAGCTTCCATGCGAACTCAAGTGGGCTCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 550 CTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGGCTGTGGAGAGAGCTTAACACCCC 609
QY 202 ArgValValG1yG1yGlnGlnuLAserValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGCTGGTGGGAGAGGCTCTGTGGAATCTTGGCCCTTGGCAGGTACGATCAG 669
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisSTrPValIleuThrAla 241
Db 670 TACGCAAAACAGCAGTCTGTGAGAGGAGCATCTGACACCCCACTGAGTCTTACGCGCA 729
QY 242 AlaHisCysPheAspArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCACGCTTTCAGGAACATACGATGTGTTCAACTGGAAGGTGGGCGAGGCTCAGAC 789

Db 1460 TCACCTATCTCAACTGATCTACATGTCTGGAAGCTGAGCTG 1504

RESULT 3

ABX76354
ID ABX76354 standard; DNA; 1314 BP.

AC ABX76354;

DT 02-APR-2003 (first entry)

XX Lung cancer-associated polynucleotide #218.

DE Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;
KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

XX WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Aziz N, Murray R;

XX WPI: 2003-093161/08.

XX P-PSDB; AB056625.

XX Detecting a lung cancer-associated transcript in a cell from a patient

XX for treating lung cancer, by contacting a biological sample from the

XX patient with a polynucleotide that exhibits increased or decreased

XX expression in lung cancer.

XX Claim 22; Page 353; 453bp; English.

XX The invention relates to a method for detecting a lung cancer-associated
XX transcript in a cell from a patient, comprising contacting a biological
XX sample from the patient with a polynucleotide that selectively hybridizes
XX to a sequence that is at least 80 % identical to a gene that exhibits
XX increased or decreased expression in lung cancer samples. Lung cancer-
XX associated polynucleotides and polypeptides are used for identifying a
XX compound that modulates a lung cancer-associated polypeptide, for
XX inhibiting proliferation of a lung cancer-associated cell to treat lung
XX cancer in a patient and for treating a mammal having lung cancer by
XX administering a modulatory compound identified. The methods are useful
XX for treating lung cancer, such as small cell lung cancer, non-small cell
XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,
XX emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
XX hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
XX bronchiectasis. The genes, polynucleotides and polypeptides are useful
XX for diagnostic purposes and as targets for screening for therapeutic
XX compounds that modulate lung cancer, such as antibodies. Sequences
XX ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the
XX invention

XX Sequence 1314 BP; 297 A; 369 C; 373 G; 275 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3,3e-207 Length: 1314
Score: 2337.00 Matches: 434

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 8 Gaps: 0

US-10-803-530-2 (1-435) x ABX76354 (1-1314)

QY	2	AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg	21
DB	10	GATCTGACAGTATTCACCTCTGACAGCTCGATGTCAAACCCCTGCGAACCCTG	69
QY	22	IlleProMetGlnThrPheArgIysValGlyIleProIleIleIleIleLeuSerLeu	41
DB	70	ATCCCCATGAGACCTTCAGAAAGTGGGGATCCCATCATCATAGACACTAGCCTG	129
QY	42	AlaSerIleIleIleValIleValIleLeuIleIleValIleLeuAspIleIleIle	61
DB	130	GCGAGTATCATATGTTGGTGTCTCATCAAGGTGATTTCTGATTAATCTACTCTC	189
QY	62	CysGlyGlnProLeuHisPheIleProArgIysGlnIleCysAspGlyGlnLeuAspCys	81
DB	190	TGCGGCGAGCTCTCCACTTCATCCAGAGAGAGCTGTGTGACGAGAGCTGACTGT	249
QY	82	ProLeuGlyGlnAspGlnIleHisCysValIysSerPheProGlnIleProAlaValAla	101
DB	250	CCCTTGGGGAGGAGCAGAGACACTGTGTCAAGACTTCCCGAAGGGCTGAGTGCA	309
QY	102	ValArgLeuSerIysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp	121
DB	310	GTCGCGCTCTCCAAAGGACCGATCCACACTGCAAGGTGTGAGCTGGCCACAGGAACTGG	369
QY	122	PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet	141
DB	370	TTCCTGCTGTTCGACACTTCACAGAGCTTCCTGTGAGAGCTGTAGGCAATG	429
QY	142	GlyTyrSerSerIysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp	161
DB	430	GGTACAGACAGCAAAACCACTTCAGAGCTGTGAGATTTGGCTTGACAGCATCGAT	489
QY	162	ValValGlnIleThrGlnAsnSerGlnIleLeuArgMetArgAsnSerGlyProCys	181
DB	490	GTTCTTCAAAATCAGAGAAACAGCAGAGAGCTTTCGATTCGCAACTCAAGTGGCCCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuThrPro	201
DB	550	CTCAGAGCTCTCGTCTCCGCTGCACTGTCTTCCTGTGGAAAGCTGAAGACCCCC	609
QY	202	ArgValValGlyGlyGlnIleAlaSerValAspSerTrpProTrpGlnValSerIleGln	221
DB	610	CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTCTTGCTTGACAGTACAGCATCAG	669
QY	222	TyrAspIysGlnHisValCysGlyIleSerIleLeuAspProHisIleTrpValLeuThrAla	241
DB	670	TACACAAACAGCAGCTGTGTGAGGAGCATCTCGAACCCCTGCTCTTCAACGCA	729
QY	242	AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValArgAlaGlySerAsp	261
DB	730	GCCACGTCTTCAGAAACATACCGATGTCTCACTGGAAGGTGGGAGGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaIysIleIleIleIleGlnPheAsnPro	281
DB	790	AAATGGGAGCTTCATCCATCTGCTGTGGCCAAAGATCATCATCATTAATTCAACCC	849
QY	282	MetTrpProIysAspAsnAspIleAlaLeuMetIysLeuGlnPheProLeuThrPheSer	301
DB	850	ATGTACCCCAAAAGCAATGACATGCGCCCTCATGTGAAGTGCAGTTCACATCTTCTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGlnIleLeuThrProAlaThrPro	321
DB	910	GGCACAGTACAGGCCCATCTGTCTCTCTTTTATATAGAGACTCATCCAGCAACCCA	969
QY	322	LeuTrpIleIleGlyTyrGlyPheThrIysGlnAsnGlyIysMetSerAspIleLeu	341

XX Serine protease; human; HUPM-6; cell proliferation; cancer;
 KW immune disorder; inflammation; therapy; ss.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 200..1507
 FT /tag= a

XX W0936550-A2.

XX 22-JUL-1999.

XX 12-JAN-1999; 99WO-US000655.

XX 16-JAN-1998; 98US-00008271.

XX (INCY-) INCYTE PHARM INC.

XX Bandman O, Hillman JL, Yue H, Guegler KJ, Corley NC, Tang YT;
 PI Shah P;
 DR WPI, 1999-430616/36.

XX P-PSDB; AAY06437.

XX Novel human protease molecules useful in the treatment of developmental
 PT disorders and/or cancers.

XX Claim 8; Page 86-87; 90pp; English.

XX This nucleotide sequence codes for HUPM-6 (see AAY06437), a novel human
 CC protease. HUPM-6 cDNA was initially identified in Incyte Clone 1337018
 CC from the colon cDNA library COLNORT13 using a computer search for amino
 CC acid sequence alignments. The present sequence is a consensus sequence
 CC derived from overlapping and/or extended nucleic acid sequences: Incyte
 CC clones 1271725 (TESTTUT02), 1337018, 586982 and 588598 (UTRSNOT01). A
 CC fragment comprising nucleotides 900-949 of the present sequence can be
 CC used for hybridisation. This sequence encompasses an active site residue.
 CC Northern analysis shows expression of HUPM-6 in gastrointestinal, and male
 CC an female reproductive cDNA libraries. Approximately 65% of these
 CC libraries are associated with neoplastic disorders, and 22% with the
 CC immune response. The invention provides 12 new human proteases, i.e. HUPM
 CC -1 to -12 (see AAY06432-43), and the polynucleotides encoding them (see
 CC AAX87149-60). Also provided are vectors, host cells and methods for
 CC producing HUPM polypeptides, as well as agonists and antagonists of HUPM.
 CC Methods for treating or preventing cell proliferative disorders and
 CC immune disorders using HUPM or HUPM antagonists are claimed

XX Sequence 2038 BP; 462 A; 591 C; 569 G; 416 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 4.76e-207 Length: 2038

XX Score: 2338.00 Matches: 434

XX Percent Similarity: 100.00% Conservative: 1

XX Best Local Similarity: 99.77% Mismatches: 0

XX Query Match: 99.83% Indels: 0

XX DB: 2 Gaps: 0

US-10-803-530-2 (1-435) x AAX87154 (1-2038)

QY 1 MetAspProAspSerAspGlnProLeuAAsenSerLeuAspValIyeProLeuAArgIysPro 20
 Db 200 ATGTGATCTGAGAGTATCACTCTGAAACAGCTCGATGTAAACCCCTGGCAAAACC 259
 QY 21 ArgGlieProMetGluThrPheArgIysValIglIleProIleIleIleAlaLeuLeuSer 40
 Db 260 CGATTCCTCATGAGAGCTTTCAGAAAGGTGGGAGATCCCATCATCATGACACTAGAC 319
 QY 41 LeuAlaSerIleIleIleValValIleuIleIysValIleLeuAspIleTyTyrPhe 60
 Db 320 CTGGGAGATCATCATTTGTGTGTCTCATCAAGGTGATTCGGATTAATACTACTTTC 379

QY 61 LeuCyGgIyGlnProLeuHiIspHeIleProArgIyGlnLeuCyAspGlyGluLeuAsp 80
 Db 380 CTCTGGGGGCGAGCTCTTCACATTCATCCCGAAGAAAGAGCTGTGTGACGGAGCTGGAC 439
 QY 81 CysProLeuGlyGluAspGluGlnHisCysValIlysSerPheProGluGlyProAlaVal 100
 Db 440 TGTCTCTGGGGGAGAGACAGAGACATGTGTCAAGACCTTCCCGAAGGGCTGGCACTG 499
 QY 101 AlaValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsn 120
 Db 500 GCAGTCCGCTCTCCAAAGAGCATTCACACTGACAGGTGCTGACTCGGCCACAGGAAAC 559
 QY 121 TrpPheSerAlaCysPheAspAsnThrThrGluAlaLeuAlaGluThrAlaCysArgGln 140
 Db 560 TGTCTCTGCTGTGCTGTTCGACAACTTCACAAAGCTTCGCTGAGACAGCTGTAGGAG 619
 QY 141 MetGlyTySerSerIysProThrPheArgAlaValIleGlyProAspGluAspLeu 160
 Db 620 ATGGGCTACAGCAGCAAAACCATTCAGACTGTGGAGATTGGCCACAGACAGGATCTG 679
 QY 161 AspValValGluIleThrGluAsnSerGlnIleuLeuArgMetArgAsnSerSerGlyPro 180
 Db 680 GATGTGTGTGAATCAGAAAAACAGCAGAGCTTCGCAATGCCGAATCAAGTGGGCC 739
 QY 181 CysLeuSerGlySerLeuValSerLeuHisCysIleuAlaCysGlyIysSerLeuIysThr 200
 Db 740 TGTCTCTAGAGCTCTGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 799
 QY 201 ProArgValValIglIyGlnGluAlaSerValAspSerTrpProTyrGlnValSerIle 220
 Db 800 CCCCCTGTGTGGGTGGGGAGAGAGCTCTGTGTGATTCCTTGCTGGCTGGCTGGCTGAG 859
 QY 221 GlnTyrAspIysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
 Db 860 CAGTACGCAAAACAGCAGCTGTGGAGGAGATCTCGAACCCCACTGGCTGCTCAGC 919
 QY 241 AlaAlaHisCysPheAspGlySerHisThrAspValPheAsnTrpIysValArgAlaGlySer 260
 Db 920 GCACCCCATCTGCTTCAGAGAAACATACCATGTGTTCATCTGAAAGGTGGCGGAGGCTCA 979
 QY 261 AspIysLeuGlySerPheProSerLeuAlaValAlaIlysIleIleIleIleGluPheAsn 280
 Db 980 GACAACTGGGAGCTTCCATCCATCCATGCTGGCTGGCCAAAGTATCATCATTAATTCAAC 1039
 QY 281 ProMetTyrProIysAspAsnAspIleAlaLeuMetIysLeuGlnPheProLeuThrPhe 300
 Db 1040 CCCATGTACCCCAAAAGCAATGACATCGCCCTCATGAAGCTGACATGCCACTCACTTTC 1099
 QY 301 SerGlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThr 320
 Db 1100 TCAGGACATGACGGCCCATCTGTCTGCTCTTTGATGAGAGCTCATCTCACCCACC 1159
 QY 321 ProLeuTrpIleIleGlyTyrGlyPheThrIysGlnAsnGlyGlyIysMetSerAspIle 340
 Db 1160 CCACTGTGATCTATGATGGGCTTTTCAGAAAGAAATGAGGGGAAGTGTCTGACATA 1219
 QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
 Db 1220 CTGCTGACAGGCTGATCCAGGTCAATGACACACAGGTGAAATGACAGACATCTGATC 1279
 QY 361 GlnGlyGluValThrGlnIysMetMetCysAlaGlyIleProGluGlnGlyValAspThr 380
 Db 1280 CAGGGGGAAGTCCAGAGAAATGATGTGTGACAGCATCCCGAAGGGGGGTGTGACACC 1339
 QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
 Db 1340 TGCAGAGGTGACAGGTGGGCCCTGATGTAACATCTGACCAATGTCAGTGGCATGTGGTGGC 1399
 QY 401 IleValSerTrpGlyTyGlyCysGlyGlyProSerThrProGlyValTyThrIleVal 420
 Db 1400 ATGTTAGCTGGGCTATGTGCTGCGGGGGCCGAGACCCCAAGAGATTAACCAAGGTC 1459
 QY 421 SerAlaTyLeuAsnTrpIleTyAsnValTrpIysAlaGluLeu 435

4 102e

XX 04-JUL-2000; 2000WO-EP006211.
XX PF 12-JUL-1999; 99EP-00113448.
XX (MERE) MERCK PATENT GMBH.
XX Suendermann B, Hofmann U, Matzku S, Wilbert O;
XX WPI: 2001-147177/15.
XX P-PSDB; AAY72558.
XX
XX New extracellular serine protease Seripancrin, useful for treating
XX cancer, osteoporosis, arthritis, chronic obstructive pulmonary disease,
XX diabetes, inflammatory disorders, stroke, angiogenesis and aberrant wound
XX healing.
XX
XX Claim 5; Page 37-39; 45pp; English.
XX
XX The present invention relates to seripancrin polynucleotides, and
XX polypeptides encoded by them. Seripancrin are members of serine protease
XX family. This protein contains a transmembrane domain, a low density
XX lipoprotein (LDL) domain, protease domain and a scavenger receptor
XX cysteine-rich (SRCR) domain. The LDL and SRCR domains help to define the
XX specificity of seripancrin's intra and intermolecular interactions. The
XX polynucleotides and polypeptides of the invention are useful for treating
XX and diagnosing diseases such as arthritis, chronic obstructive pulmonary
XX disorder (COPD), cancer, osteoporosis, aberrant wound healing,
XX angiogenesis, inflammatory disorders, diabetes, stroke and cardiovascular
XX diseases. Seripancrin genes are useful in chromosome localization
XX studies, as tools for tissue expression studies and also in gene therapy.
XX The polypeptides of the invention are used for identifying agonists and
XX antagonists useful for treating conditions associated with seripancrin
XX imbalance. These polypeptides are also useful as vaccines. The present
XX sequence is a cDNA coding for seripancrin protein. The seripancrin gene
XX is located on human chromosome 11q22-q23
XX
XX Sequence 1305 BP; 293 A; 368 C; 372 G; 272 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.12e-207 Length: 1305
XX Score: 2342.00 Matches: 435
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 4 Gaps: 0
XX
XX US-10-803-530-2 (1-435) x AAD02556 (1-1305)
XX
XX 1 MetAspProAspSerAspGlnProLeuHisSerLeuAspValLysProLeuArgLysPro 20
XX 1 ATGGATCTGACAGATCAACCTCTGAACAGCTCGATGTCMAACCCCTGGGCAAAACC 60
XX
XX 21 ArgIleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuDeuSer 40
XX 61 CGTATCCCAAGAGACCTTCAGAAAGGTGGGAGATCCCATCATATAGCACTACTAGCC 120
XX
XX 41 LeuAlaSerIleIleIleValValValLeuIleLysValIleLeuAspLysTyrTrpHe 60
XX 121 CTGGGAGTATCATCATGTGTGTCTCTCATCAAGTGTGATTCGATAATACTACTTC 180
XX
XX 61 LeuGlyGlyGlnProLeuHisPheIleProArgLysGlnLeuGlyAspGlyGlnLeuAsp 80
XX 181 CTCTGGGGGAGCCCTCTCCACTTCATCCCGAAGAGCAAGCTGTGTGACGGAGAGCTGAGAC 240
XX
XX 81 CysProLeuGlyGlnLysAspGlyGlnLysIleCysValLysSerPheProGlyGlyProAlaVal 100
XX 241 TGTCTCTTGGGGGAG 300
XX
XX 101 AlaValAlaGLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsn 120
XX 301 GCAGTCCGCTCTCCCAAG 360

QY 121 TrpPheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGln 140
DB 361 TGGTTCCTGCTGCTTTTTCGAACTTCACAGAACTCTCGCTGAGAGAGCTGTAGAGCAG 420
QY 141 MetGlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeu 160
DB 421 ATGGGCTTACAGACAGCAAAACCATTTGAGAGCTGTGAGATGGCCAGACAGAGATCTG 480
QY 161 AspValAlaGluIleThrGlnLysSerGlnGluLeuArgMetLysAsnSerSerGlyPro 180
DB 481 GATGTGTGAATCAGAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThr 200
DB 541 TGTCTCTCAGGCTCTCGGTCTCTGCTCAGCTCTTCTGCGTGTGGAGAGAGCTGAGAGCC 600
QY 201 ProArgValValGlyGlyGlnLysValSerValAspSerTrpProTrpGlnValSerIle 220
DB 601 CCCCCTGTGGGTGGGGAG 660
QY 221 GlnTyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTyrValLeuThr 240
DB 661 CAGTACAGCAAAACAG 720
QY 241 AlaAlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySer 260
DB 721 GCAGCCCACTGCTCTCAGAGAAACATACCGATGTCTTCAACTGGAAGTCCGGGAGGCTCA 780
QY 261 AspLysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsn 280
DB 781 GACAACTGGGAG 840
QY 281 ProMetLysProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrHe 300
DB 841 CCCATGTAACCCCAAGCAAG 900
QY 301 SerGlyThrValArgProIleCysLeuProPheAspGlnGluLeuThrProAlaThr 320
DB 901 TCAGGACAGTCAAG 960
QY 321 ProLeuTrpIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
DB 961 CCACCTGTGATCATTTGAG 1020
QY 341 LeuLeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
DB 1021 CTGCTGAGGGGTGATTCAGAGTCAATGACAGACAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 361 GlnGlyGlnValThrGlnLysMetMetCysAlaGlyIleProGluGlyValAspThr 380
DB 1081 CAGGGGAGAGTCAAG 1140
QY 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGly 400
DB 1141 TGGCAGGGGTGACAGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 401 IleValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysVal 420
DB 1201 ATCGTTAGCTGGGGCTATGCTGGGGGGCCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 421 SerAlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
DB 1261 TCAAGCTATCTCAAGTGAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1305
XX
XX RESULT 2
XX ID AAX87154 standard; cDNA; 2038 BP.
XX AC AAX87154;
XX XX 27-SEP-1999 (first entry)
XX DE Human protease HUPM-6 cDNA.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 6, 2005, 18:45:35 / Search time 680 Seconds

(without alignments)
4263.443 Million cell updates/sec

Title: US-10-803-530-2

Perfect score: 2342
Sequence: 1 MPPDDQPLNSLDVXPKRKP.....VYTKVSAYLWNIYVWYKAEI 435

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus.p2n.model -DEV=x1p
-O/cgr2.1/USPTO_spool_p/US10803530/runat_05122005_083215_2111/app_query.fasta_1.583
-DB=N_Geneseq -QFMT=fastap -SUFFIX=ring -MINMATCH=0 -1 -LOOEXT=0 -LOOEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdt -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10803530 @CGN_1_1-1096 @runat_05122005_083215_2111 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUDERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342	100.0	1305	4	AAD02556 Human ser
2	2338	99.8	2038	2	AAx87154 Human pro
3	2337	99.8	1314	8	ABx76354 Lung can
4	2337	99.8	1314	10	ADB80524 Ovarian c

5	2337	99.8	1314	11	ADN39512 Cancer/an
6	2337	99.8	1314	11	ADN39661 Cancer/an
7	2337	99.8	1314	11	ADN39472 Cancer/an
8	2337	99.8	1314	11	ADN39815 Cancer/an
9	2337	99.8	1314	11	ADN39460 Cancer/an
10	2337	99.8	1314	13	ADU25739 CDNA enco
11	2337	99.8	2081	6	AA172976 CVA8 pref
12	2337	99.8	2104	12	ADN04463 Antispori
13	2337	99.8	2104	14	ADW04403 Human Ova
14	2337	99.8	2104	14	AE554844 Human TAT
15	2337	99.8	2307	6	ABF76529 Human cer
16	2337	99.8	2307	8	ACF12937 Human cer
17	2335	99.7	2121	6	AA164284 Human ser
18	2333	99.6	2590	13	ABZ35523 Human gen
19	2329	99.4	1479	6	ABD32787 Human can
20	2329	99.4	2165	6	ABZ35523 Human ser
21	2328	99.4	2627	13	ABD32788 Human gen
22	2324	99.2	2070	3	AA290471 Cancer can
23	2324	99.2	2079	4	AA290471 Cancer sp
24	2324	99.2	2079	6	AA172975 Human tra
25	2324	99.2	2079	9	ACF36041 Human tra
26	2324	99.2	2079	10	ADAE48026 DNA enco
27	2324	99.2	2079	10	ADAE48026 DNA enco
28	2324	99.2	2079	10	ADAE48026 DNA enco
29	2324	99.2	2079	12	ABG08800 Human cel
30	2324	99.2	2079	12	ABG08800 Transmem
31	2324	99.2	2079	12	ADJ46931 Human tra
32	2324	99.2	2079	13	ADAE48219 Human tra
33	2324	99.2	2079	14	ABE35225 Human typ
34	2319	99.0	2137	10	AD110372 Human mem
35	2319	99.0	2137	12	ADJ46896 Human cel
36	2297.5	99.1	2063	3	AAJ37099 Human PRO
37	2297.5	99.1	2063	4	AAJ37099 Human PRO
38	2297.5	99.1	2063	4	AAJ37099 Human PRO
39	2297.5	99.1	2063	4	AAJ37099 Human PRO
40	2297.5	99.1	2063	6	AAJ37099 Human PRO
41	2297.5	99.1	2063	6	AAJ37099 Human PRO
42	2297.5	99.1	2063	8	ABK10930 Human CDN
43	2297.5	99.1	2063	8	ABK10930 Human CDN
44	2297.5	99.1	2063	8	ABK10930 Human CDN
45	2297.5	99.1	2063	8	ABK10930 Human CDN

ALIGNMENTS

RESULT 1	
AAD02556	standard; cDNA, 1305 BP.
ID	AAD02556
AC	AAD02556
XX	
DT	02-MAY-2001 (first entry)
XX	
DE	Human seripancrin cDNA.
XX	
KW	Human; seripancrin; serine protease; chromosome 11q22-q23; therapy;
KW	arthritis; chronic obstructive pulmonary disorder; COPD; cancer;
KW	osteoporosis; aberrant wound healing; angiogenesis; diabetes;
KW	inflammatory disorder; stroke; cardiovascular disease; gene therapy;
XX	vaccine; cytosolic; cerebroprotective; vulnerary; osteopathic; ss.
OS	Homo sapiens.
XX	
PH	
FT	Key
FT	CDS
FT	1..1305
FT	/tag= a
FT	/product= "Human seripancrin protein"
FT	/note= "The coding region does not include stop codon"
XX	
XX	MO200104141-A2.
XX	18-JAN-2001.

```

Db 1304 CTGCAGGCGTCAGTCCAGATGACACACACGGTGCAATGCAGACGATGCGTACCAG 1363
Qy 362 G|YGIuValThrGluValMetMetCysAlaGlyIleProGluGlyValAspThrCys 381
Db 1364 GGGGAAGTCACCGAAGATGATGTGTGACAGGCATCCGGAAGGGGTGTGGACACCTGC 1423
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1424 CAGGCTGACAGTGGTGGGCCCCCTGATGTACCAATCTGACAGTGGCATGTGTGGGCATC 1483
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1484 GTTAGCTGGGGCTATGGCTGGGGGGCCGAGCACCCGAGAGTATACACCAAGGTCTCA 1543
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1544 GCCTATCTCAACTGAGTCTACAAATGTCTGGAAGGCTGAGCTG 1585

```

Search completed: December 7, 2005, 01:13:14
 Job time : 932 secs

Db 1184 GGCAAGTCAGGCCCATCTGTCTGCTCTTTGATGAGAGCTCACTCCAGCCACCCCA 1243
QY 322 LeuTriP11e11eg1yTriPglyPheThrLySGlnAsnGlyLyMetSerAsp11e1eu 341
Db 1244 CTCTGGATCATTTGATGGGGCTTTTACGAAGCAGAAATGAGGAGAAATCTGACATCTG 1303
QY 342 LeuGln1AspSerValGlnVal11eAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1304 CTGACGGGCTGATGCTCAGGTATTCAGACACAGGTGCAATTCAGACGATGCTCAACAG 1363
QY 362 GlyGlnVal1ThrGlnLyMetMetCysAlaGly11eProGlnGlyGlyVal1AspThrCys 381
Db 1364 GGGGAAGTCACGAGAAATATGCTGTCCAGGATCCCGAAGGGGGTGGACACTGC 1423
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrglnSerAspGlnTriP11eVal1Gly11e 401
Db 1424 CAGGTGACAGTGTGGGCCCCGTGATGACCAATCTGACCACTGGCATGTGGGGCATC 1483
QY 402 ValSerTriPglyTyrglyCysGlyGlyProSerThrProGlyVal1TyrrThrLyValSer 421
Db 1484 GTTAGCTGGGGCTTGGCTGGGGGGCCCCGAGCACCCGAGAGATATCACCAAGGTCTCA 1543
QY 422 AlaTyrrLeuAsnTriP11eTyrrAsnVal1TriPlyAlaGly11e 435
Db 1544 GCCTATCTCAACTGGATCTACAAATGTCTGGAGGCTGAGCTG 1585

RESULT 15

US-10-171-311-217
Sequence 217, Application US/10171311
Publication No. US20030087270A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Chen, Yan
APPLICANT: Zhao, Xumei
APPLICANT: Monahan, John
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Glatt, Karen
APPLICANT: Gannavarapu, Manjula
APPLICANT: Hoerh, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171,311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 217
LENGTH: 2307
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: m18c_feature
LOCATION: 1691, 1698, 1705, 1708, 1709, 1713, 1717, 1720, 1724, 1728,
LOCATION: 1731, 1741, 1746, 1748, 1755, 1770, 1774, 1791, 1802, 1821,
LOCATION: 1838, 1856, 1859, 1864, 1908, 1959, 1997, 2012, 2038, 2143
OTHER INFORMATION: n = A,T,C or G
US-10-171-311-217

Alignment Scores:

Pred. No.: 1.8e-282 Length: 2307
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 5 Gaps: 0

US-10-803-530-2 (1-435) x US-10-171-311-217 (1-2307)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspVal1LyProLeuArgLyProArg 21
Db 284 GATCTGACAGATGATTCACCTTCTGAAAGCCCTCGATGTCAAACCCCTCGCAACCCCGT 343
QY 22 IleProMetGlnThrPheArgLyValGly11ePro11e11e11e11e11e11e11e11e11e 41
Db 344 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATATATAGCACTAGAGCTG 403
QY 42 AlaSer11e11e11eVal1Val1leu11e11e11e11e11e11e11e11e11e11e11e11e11e 61
Db 404 GCGAGTATCATATGTGGTGTCTCTCATCAAGGTGATTCGATTAATATCTCTCTC 463
QY 62 CysGlyGlnProLeu11ePhe11eProArgLyGlnLeuCysAspGlyGlyLeuAspCys 81
Db 464 TGGGGAGCCCTCTCACCTTCATCCCGAAGAGCAGTGTGTGACGGAGACTGACTGT 523
QY 82 ProLeuGlyGlyAspGlyGly11eCysVal1LySerPheProGlyGlyProAlaValAla 101
Db 524 CCTTGGGGGAGAGAGAGAGACTGTCTCAAGACTTCCCGAAGGGCTGTGAGTGGCA 583
QY 102 ValArgLeuSerLyAspArgSerThreuglnVal1leuAspSerAlaThrGlyAsnTriP 121
Db 584 GTCCGCTCTCCAGAGCCGATCCACACTGAGGTGCTGAGCTGGCCACAGGAACTGG 643
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAla1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e1e 141
Db 644 TTCTTGCTGCTTTTCGACAACTTCACAGAACTCTCTGAGACGCTGTGAGCACTGTGAGCAATG 703
QY 142 GlyTyrrSerSerLyProThrPheArgAlaValGly11eGlyProAspGlnAspLeuAsp 161
Db 704 GGTACACAGACAAACCACTTTCAGAGCTGTGAGATTTGGCCAGACAGATCTGAT 763
QY 162 ValValGly11eThrGlnAsnSerGlnGlyLeuArgMetArgAsnSerGlyProCys 181
Db 764 GTTGTGAATCACAGAAACAGACAGAGCTTCGATGCGAATCGGAGCTCAAGTGGGCTGT 823
QY 182 LeuSerGlySerLeuValSerLeu11eCysLeuAlaCysGlyLySerLeuLyThrPro 201
Db 824 CTCTCAGGCTCTCCGTGCTCTCTCTCAGCTGTCTTCTGTGGAGAGACCTGAAAGACCC 883
QY 202 ArgValValGlyGlyGlnAlaSerValAspSerTriProTriP11eGln 221
Db 884 CGTGTGGGTGGGGAGAGAGCTCTGTGATTTCTTGCTTGGCAGGTGACATCCAG 943
QY 222 TyrrAspLySGln11eVal1CysGlyLySer11eLeuAspProhiSTrPVal1LeuThra 241
Db 944 TACGACAAACAGCAGTCTGTGAGGAGCATCTGACCCCACTGGGTCTCACGGGCA 1003
QY 242 Ala11eCysPheArgLySh1SThrAspVal1PheAsnTriPlyVal1ArgAlaGlySerAsp 261
Db 1004 GCCACGTCTTCAGAGAAACATCCGATGTTCACACTGAAAGGTGGGGGCTCAAGAC 1063
QY 262 LysLeuGlySerPheProSerLeuAlaVal1AlaLy11e11e11e11e11e11e11e11e11e11e 281
Db 1064 AAACGTGGGAGCTTCCATCTCTGGCTGTGGCCAAAGATCATCATTTGAATTCACACCC 1123
QY 282 MetTyrrProLyAspAsnAsp11eAla1e 301
Db 1124 ATGTACCCCAAAAGACATGACATGCGCTCATATMACTGAGTCCACACTCATCTTCTCA 1183
QY 302 GlyThrValArgPro11eCysLeuProPheAspGlyGlyLeuThrProAlaThrPro 321
Db 1184 GGCAAGTCAGGCCCATCTGTCTGCTCTTTGATGAGAGCTCACTCCAGCCACCCCA 1243
QY 322 LeuTriP11e11eg1yTriPglyPheThrLySGlnAsnGlyLyMetSerAsp11e1eu 341
Db 1244 CTCTGGATCATTTGATGGGGCTTTTACGAAGCAGAAATGAGGAGAAATCTGACATCTG 1303
QY 342 LeuGln1AspSerValGlnVal11eAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361

Db 1202 CTCTGGATCATTCGATGGGGCTTTACGACGACGAGTGGAGGAGTCTGACACTA CTG 1261
Qy 342 LeuGlnIAspSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1262 CTGAGGGGTAGTCAGGTATTCAGACACAGCGGTGCATGCAATGCAGATGCCGTAACCG 1321
Qy 362 G1yGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1322 GGGGAACTCACCGAAGATATGTGTCCAGGCATCCCGAAGGGGGGTGTGACACCTGC 1381
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheIleValGlyIle 401
Db 1382 CAGGCTGACAGTGGTGGGCCCTGATGTACCAATCTGACACAGTGCAATGTGGGCACTC 1441
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1442 GTTAGCTGGGGCTATGCTGGGGGGCCGAGCACCCGAGATATACCAAGGCTTCA 1501
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTyrLysAlaGluLeu 435
Db 1502 GCCTATCTCAACTGATCTACAAATGTCTGGAAGGCTGAGCTG 1543
RESULT 14
US-10-097-340-317
Sequence 317, Application US/10097340
Publication No. US20030087250A1
GENERAL INFORMATION:
APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVARAPU
APPLICANT: Sebastian HOERSCH
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVARS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VEIRY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, JR.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 317
LENGTH: 2307
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(2307)
OTHER INFORMATION: n = A,T,C or G

US-10-097-340-317
Alignment Scores:
Pred. No.: 1,8e-282 Length: 2307
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 5 Gaps: 0
US-10-803-530-2 (1-435) x US-10-097-340-317 (1-2307)
Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIleAspProLeuAspArg 21
Db 284 GATCCTGACAGTATCAACCTCTGAACAGCCTCGATGTCAAAACCCCTGCCGAACCCCGT 343
Qy 22 IleProMetGluThrPheAspGlyValGlyIleProIleIleIleAlaLeuSerLeu 41
Db 344 ATCCCATGGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCACTAAGGCTG 403
Qy 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleVal 61
Db 404 GCGAGTATCATATTGGTTGTTCTCTCATCAAGTGATTTGGATTAATCTACTTCTC 463
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 464 TGGGGGAGGCTCTCCACTTCATCCGAGGAGCAGCTGTGTGACGAGAGCTGACTGT 523
Qy 82 ProLeuGlyIleAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 524 CCTTGGGGGAGGAGCGAGGACCTGTGTCAAGGCTTCCCGGAAGGCTGCAAGTGCA 583
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleAspSerAlaThrGlyAsnTrp 121
Db 584 GTCCGCTCTCAAGACCCATCCATCCAGAGTGCTGAGCTGGCCACAGGAACTGG 643
Qy 122 PheSerAlaCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 644 TTCCTGCTGTTTGCACAACTTCACAGAGCTCTCGTGAGACAGCTGTAGGACATG 703
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 704 GGTACAGACAGCAACCACTTTCAGAGCTGTGAGATGTGCCACAGCATCTGAT 763
Qy 162 ValValGluIleThrGluAsnSerGlnGluLeuArgPheAlaGlyAsnSerSerGlyProCys 181
Db 764 GTTGTGAAATCAGAAACAGACGAGAGCTTGCAGTGGAACTCAAGTGGGCTGT 823
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro 201
Db 824 CTTCAGGCTCCTCGTGTCTCCCTGCACTGTCTTGCCTGTGGAAAGGCTTAAGACCCC 883
Qy 202 ArgValValGlyGlyGluGluAlaSerValAspSerTrpProIleGlnValSerIleGln 221
Db 884 CGTGGGTGGGTGGGAGAGAGGCTCTGTGATTCTTGCTGGCAGTGCAGCATTCAG 943
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisIleThrValLeuThrAla 241
Db 944 TACGCAAAACAGACGCTGTGGAGGAGCATCTCGAGCCCACTGGGTCTTCACGGCA 1003
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 1004 GCCACCTCTTCAGGAACATACGATGTGTTCACATGGAAGTGGGGGAGGCTCAAC 1063
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaIleIleIleIleIleGluPheAsnPro 281
Db 1064 AAACGGGAGCTTCCATCCCTGGCTGTGGCAAGATCATCATATTGAATTCAACCCC 1123
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1124 ATGTACCCCAAGAAAGATGACATCCCTCATGAGCTGCAAGTGTCCACTCATCTTCTCA 1183
Qy 302 GlyThrValArgProIleCysLeuProPheAspGluGluLeuThrProAlaThrPro 321

QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
Db 902 TACACAAAGAGCAGCTGTGTGAGGAGCATCTCGAACCCCACTGGCTCTCACGGCA 961
QY 242 AlaHisCysPheArgLysHisTrpAspValPheLeuTrpLysValIleGlnAglySerAsp 261
Db 962 GCCACCTGCTTCAGGAAACATACGATGCTTCACTGAAAGGCGGGAGGCTCGAC 1021
QY 262 LysLeuGlySerPheProSerIleuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 1022 AAACCTGGGAGCTTCCCATCCCTGCTGTGGCCAAAGTCAATCATCAATTGAATCAACCCC 1081
QY 282 MetTrpTrpLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACAATGACATGCCCCATGAAAGTTCACATCCACTCTCTCA 1141
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1142 GGCACAGTCAGGCCCATCTGTCTGCTCTTGTGATGAGAGCTCACTCCAGCCACCCCA 1201
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341
Db 1202 CTCGTGATCATTTGGATGGGCTTTAGCAAGACAGATGAGAGGATGTGTGACATACCTG 1261
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrpGln 361
Db 1262 CTCACAGCGTCAGCTCCAGCTCATTTGACACACACGCGTGCATGACAGATCGCTACAG 1321
QY 362 GlyGlnValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThrCys 381
Db 1322 GGGGAAGTCACCGAAGATGATGTGTGACAGGATCCCGAAGGGGTGTGGACACTCGC 1381
QY 382 GlnGlyAspSerGlyGlyProLeuMetTrpGlnSerAspGlnTrpHisValIleGlyIle 401
Db 1382 CAGGGTGACAGTGTGGGCCCCGTGATGTACCAATCTGACACAGTGGCATGTGGGCAATC 1441
QY 402 ValSerTrpGlyTrpGlyCysGlyGlyProSerThrProGlyValTrpThrLysValSer 421
Db 1442 GTTAGCTGGGGCTATGCTGTGGGGGGCCCGACACCCCAAGAGTATACCAAGGCTCTCA 1501
QY 422 AlaTrpLeuAsnTrpIleTrpAsnValTrpLysAlaGluLeu 435
Db 1502 GCTATCTCACTGATCTACATGTCTGTGAGAGGCTGAGCTG 1543
RESULT 13
US-10-991-287-2
Sequence 2, Application US/10991287
Publication No. US20050208523A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Polakis, Paul
APPLICANT: Smith, Victoria
APPLICANT: Wood, William I.
APPLICANT: Wu, Thomas D.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE OF INVENTION: TREATMENT OF TUMOR
FILE REFERENCE: P5037R1-US
CURRENT APPLICATION NUMBER: US/10/991,287
PRIOR FILING DATE: 2004-11-17
PRIOR APPLICATION NUMBER: US 60/523, 856
NUMBER OF SEQ ID NOS: 10
SEQ ID NO 2
LENGTH: 2104
TYPE: DNA
ORGANISM: Homo sapiens
US-10-991-287-2
Alignment Scores: 1.56e-282 Length: 2104
Pred. No.: 1

Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 9 Gaps: 0
US-10-803-530-2 (1-435) x US-10-991-287-2 (1-2104)
QY 2 AspProAspSerAspGluProLeuAsnSerIleuAspValLysProLeuArgLysProArg 21
Db 242 GATCTGACAGTATCACTGATCACTGAAAGCTTCGATGTCAAAACCCCTGGCAAAACCCGT 301
QY 22 IleProMetGluThrPheArgLysValGlyIleProIleIleIleAlaLeuLeuSerLeu 41
Db 302 ATCCCATGGAACCTTCAGAAAGGTGGGATCCCATCATATATAGCATCTGAGCTTG 361
QY 42 AlaSerIleIleIleValValIleuIleLysValIleLeuAspLysTrpThrPheLeu 61
Db 362 GCGAGTATCATTTGTGGTGTCTCTCATCAAGGTGATTTCTGGATTAATTAATCACTTCTC 421
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGluLeuAspCys 81
Db 422 TCGGGCAGCTCTCTCACTTCAATCCCGAAGAGCAGCTGTGTGACGAGAGCTGACCTGT 481
QY 82 ProLeuGlyGluAspGluGluHisCysValLysSerPheProGluGlyProAlaValAla 101
Db 482 CCTTGGGGGAGAGAGAGAGACATGTGTCAAGACTTCCCGAAGGCTGTGACGTGCA 541
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 542 GTCCGCTCTCCAAAGACCGATCCACATGACAGGTGTGACTCGGCACAGGGAATCTGG 601
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluTrpAlaCysArgGlnMet 141
Db 602 TTCTCTGCTGTTTTCAGAACTTTCACAAAGCTCTCGTGGAGACGCTGTAGGACATG 661
QY 142 GlyTrpSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 662 GGCACACAGAGAAACCACTTTCAGAGCTGTGAGATTTGGCCACAGACAGATCTGGAT 721
QY 162 ValValGluIleTrpGlnAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 722 GTTGTGAATCACAGAAACAGGACGAGGCTTCGATGCGGAATCAAGTGGGCCCTGT 781
QY 182 LeuSerGlySerLeuValSerLeuHisCysIleuAlaCysGlyLysSerLeuLysThrPro 201
Db 782 CTCCTAGGCTCCCTGGTGTCTCTGACTGTCTGTGTGGAAAGCTGAAAGACCCCC 841
QY 202 ArgValValIleGlyGlyGluAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 842 CGTGGTGGTGGGGAGAGAGGCTGTGTGATTTCTTGGCTTGGCAGGTACGATCCAG 901
QY 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
Db 902 TACGACAAAGCAGCAGTGTGTGAGGAGCATCTCGAACCCCACTGGGCTCTCACGGCA 961
QY 242 AlaHisCysPheArgLysHisTrpAspValPheLeuTrpLysValIleGlnAglySerAsp 261
Db 962 GCCACCTGCTTCAGGAAACATACGATGCTTCACTGAAAGGCGGGAGGCTCGAC 1021
QY 262 LysLeuGlySerPheProSerIleuAlaValAlaLysIleIleIleGluPheAsnPro 281
Db 1022 AAACCTGGGAGCTTCCCATCCCTGCTGTGGCCAAAGTCAATCATCAATTGAATCAACCCC 1081
QY 282 MetTrpTrpLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 1082 ATGTACCCCAAGACAATGACATGCCCCATGAAAGTTCACATCCACTCTCTCA 1141
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
Db 1142 GGCACAGTCAGGCCCATCTGTCTGCTCTTGTGATGAGAGCTCACTCCAGCCACCCCA 1201
QY 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyLysMetSerAspIleLeu 341

QY 122 PheSerAlaCysPheAspAsnPhenThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 602 TTCTCTGCTGCTTTTCGACAACTTTCACAGAACTCTCGCTGAGACAGCTTGTAAGCAGATG 661
QY 142 G1YrSerSerLysProThrPheArgAlaValGluIleG1YProAspGlnAspLeuAsp 161
DB 662 GCGTACAGACAGCAAAACCACTTTCAGAGCTGTGGAGATTTGGCCAGACAGAGATCTGAT 721
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetAlaGlnSerSerG1YProCys 181
DB 722 GTTGTGTAATCACAGAAACAGCCAGAGCTTCCATGCGGAACTCAAGTGGGCCCTGT 781
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysG1YLySerSerLeuLysThrPro 201
DB 782 CTCACAGGCTCCCTGGTCTCCCTGCATGCTCTTCTGGAGAAAGCTGAAAGCCCCC 841
QY 202 ArgValValG1YGLYGLUGluAlaSerValAspSerTPPProTPRGLValSerIleGln 221
DB 842 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTTGGCTTGGCAGGTCCAGATCCAG 901
QY 222 TyrAspLysGlnHisValCysG1YGLYSerIleLeuAspProHisTPRValLeuThrAla 241
DB 902 TACGACAAACAGCAGCTGTGTGAGAGGAGCATCTGAGACCCCACTGGGCTCTCAGCGCA 961
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTPRValArgAlaG1YSerAsp 261
DB 962 GCCCACTCTTCAGAGAAACATACCATGATGTTCACATGAGAGGTCCGGGAGGCTCAGAC 1021
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGluPheAsnPro 281
DB 1022 AAACGGGAGCTTCCATCCCTGGCTGTGGCAAGATCATCATTAATTAATCAACCCC 1081
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 1082 ATGTACCCCAAGACATGACATGCCCTCATGAAGCTGCACTCCACTCACTTCTCA 1141
QY 302 G1YThrValArgProIleCysLeuBProPheAspGluGluLeuThrProAlaThrPro 321
DB 1142 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGAGATCATCTCCAGCCCA 1201
QY 322 LeuTPRleIleG1YrG1YrPheThrLysGlnAsnG1YLyLysMetSerAspIleLeu 341
DB 1202 CTCTGGATCATTTGATGGGCTTTTACGAAAGCATGAGGAGATGTCTGACACTACTG 1261
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1262 CTGCGGGGTCACTCAAGTCACTTGAACAGACACAGGTGCAATGACAGAGTGCATCAG 1321
QY 362 G1YGLYValAlaThrGluLysMetCysAlaG1YIleProGluG1YGLYValAspThrCys 381
DB 1322 GGGGAGATCACCGAAGAATGATGTGTGAGGCACTCCGGAAGGGGTGTGACACCTGC 1381
QY 382 GlnGlyAspSerG1YGLYProLeuMetTyrGlnSerAspGlnTPRHisValValIle 401
DB 1382 CAGGTGACAGGTGTGGGCCCTGTATGCCATCTGACCACTGACATGTGTGGGCAATC 1441
QY 402 ValSerTPRGLYrG1YrG1YrCysG1YGLYProSerThrProG1YValTyrThrLysValSer 421
DB 1442 GTTACTGGGGCTAAGGCTGCGGGGGCCCGAGCACCCCGAGGTATACACCAAGGTCTCA 1501
QY 422 AlaTyrLeuAsnTPRleIleTyrAsnValTPRValAspIleLeu 435
DB 1502 GCTATCTCACTGATCTACAAATGTCTGAAAGGCTGAGCTG 1543

RESULT 12
us-10-994-117-2

; Sequence 2, Application US/10994117

; Publication No. US20050164250A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi J.

; APPLICANT: Goddard, Audrey

; APPLICANT: Gurney, Austin L.

; APPLICANT: Polakis, Paul
; APPLICANT: Smith, Victoria
; APPLICANT: Wood, William I.
; APPLICANT: Wu, Thomas D.
; APPLICANT: Zhang, Dong-Xiao
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: TREATMENT OF TUMOR
; FILE REFERENCE: P503782
; CURRENT APPLICATION NUMBER: US/10/994,117
; PRIOR FILING DATE: 2004-11-19
; PRIOR APPLICATION NUMBER: US 60/523,856
; PRIOR FILING DATE: 2004-11-20
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 2
; LENGTH: 2104
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-994-117-2
Alignment Scores:
Pred. No.: 1,56e-282 Length: 2104
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: Gaps: 0
US-10-803-530-2 (1-435) x US-10-994-117-2 (1-2104)
QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValLysProLeuArgLysProArg 21
DB 242 GATCTGACAGTATCAACCTTGAACAGCTCGATGATCAAAACCCCTGGCAAAACCCCGT 301
QY 22 IleProMetGluThrPheArgLysValG1YIleProIleIleIleAlaLeuLeuSerLeu 41
DB 302 ATCCCATGAGACCTTCAGAAAGTGGGATCCCATCATCATAGCATCTAGCCTG 361
QY 42 AlaSerIleIleIleValValLeuIleLysValIleLeuAspLysTyrTPRLeu 61
DB 362 GCGAGTATCATATTGTGTGTCTCTCATCAAGGTGATTCGATTAATACTACTCTC 421
QY 62 CysG1YGLNProLeuHisPheIleProArgLysGlnLeuCysAspG1YGLYLeuAspCys 81
DB 422 TGCAGGAGCTCTCCACTTCATCCAGAGAAAGCATGTGTGAGAGAGCTGACTGT 481
QY 82 ProLeuGlyLysAspGluGluHisCysValLysSerPheProGluG1YProAlaValAla 101
DB 482 CCTTGGGGAGAGAGAGACATGTGTCAAGACTTCCCGAAGGCTGCAGTGCA 541
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValLeuAspSerAlaThrG1YAsnTPR 121
DB 542 GTCCGCTCTCAAGAGACCATTCACACTGAGGTGTGACTGGCCACAGGGAACTGG 601
QY 122 PheSerAlaCysPheAspAsnPhenThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
DB 602 TTCTCTGCTGCTTTTCGACAACTTTCACAGAACTCTCGCTGAGACAGCTTGTAAGCAGATG 661
QY 142 G1YrSerSerLysProThrPheArgAlaValGluIleG1YProAspGlnAspLeuAsp 161
DB 662 GCGTACAGACAGCAAAACCACTTTCAGAGCTGTGGAGATTTGGCCAGACAGAGATCTGAT 721
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetAlaGlnSerSerG1YProCys 181
DB 722 GTTGTGTAATCACAGAAACAGCCAGAGCTTCCATGCGGAACTCAAGTGGGCCCTGT 781
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysG1YLySerSerLeuLysThrPro 201
DB 782 CTCACAGGCTCCCTGGTCTCCCTGCATGCTCTTCTGGAGAAAGCTGAAAGCCCCC 841
QY 202 ArgValValG1YGLYGLUGluAlaSerValAspSerTPPProTPRGLValSerIleGln 221
DB 842 CGTGTGGTGGTGGGAGAGAGGCTCTGTGATTTTGGCTTGGCAGGTCCAGATCCAG 901

Query Match: 99.79% Indels: 0
 DB: 3 Gaps: 0
 US-10-803-530-2 (1-435) x US-09-851-588-7 (1-2081)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 DB 224 GATCCTGACAGTGAATCAACCTCTGAACAGCTCCGATGCAAAACCCCTGGCAACCCCGT 283
 QY 22 IleProMetGluThrPheArgIysValGlyIleProIleIleIleIleAlaLeuLeuSerLeu 41
 DB 284 ATCCCAAGAGAGACCTTCAGAAAGGTGGGATCCCAATCATCATAGCACTGAGCGCTG 343
 QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
 DB 344 GCGAGTATCATCATGTGTGTGTCTCTCATCAAGGTGATTCGGATTAATTAATTAATTCCTC 403
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
 DB 404 TGGGAGAGCTCTCCCACTTCATCCAGAGAACAGCTGTGTGACGAGAGCTGGACTGT 463
 QY 82 ProLeuGlyGluAspGluGluHisCysValIysSerPheProGluGlyProAlaValAla 101
 DB 464 CCCTTGGGGAG 523
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
 DB 524 GTCCGCTCTCCAG 583
 QY 122 PheSerIleCysPheAspAsnPheThrGluAlaLeuAlaGluThrAlaCysArgGlnMet 141
 DB 584 TTCTGCTGCTGTTCCAGCAACTTCACAGAAAGCTCTGCTGAGACAGCTGTAGAGAGAG 643
 QY 142 GlyTyrSerSerIysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
 DB 644 GGTACAG 703
 QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
 DB 704 GTTGTGAATTCACAGAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 763
 QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
 DB 764 CTCTCAGGCTCCCTGTGCTCCCTGCACTGTCTTGCCTGTGGAGAGAGAGAGAGAGAG 823
 QY 202 ArgValIleGlyGluGluGluIleSerValAspSerTrpProTrpGlnValSerIleGln 221
 DB 824 CGT 883
 QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValLeuThrAla 241
 DB 884 TAGCAACAACAGACCTCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 943
 QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpIysValArgAlaGlySerAsp 261
 DB 944 GCCCATGTGTTACAGAAACATACCGATGTGTTCACTGGAAGTGGGGAGAGAGAGAGAG 1003
 QY 262 LysLeuGlySerPheProSerLeuAlaValAlaIysIleIleIleIleGluPheAsnPro 281
 DB 1004 AAACCTGGAGAGCTTCCATCCCTGCTGTGGCTGTGGCAAGATCATCATGAAATTCACACCC 1063
 QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
 DB 1064 ATGTACCCCAAG 1123
 QY 302 GlyThrValArgProIleCysLeuProPhePheAspGluGluLeuThrProAlaThrPro 321
 DB 1124 GGCACAGTACAGCCCATCTGTGCTGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1183
 QY 322 LeuTrpIleIleGlyTrpGlyPheThrIysGlnAsnGlyGlyLysMetSerAspIleLeu 341
 DB 1184 CTCTGATCATTTGATGGGGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1243

QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
 DB 1244 CTGACAGGCTCAGTCCAGGTTCATTCAGACACAGCGTGCATTCAGAGATGCTACACAG 1303
 QY 362 GlyGluValIleThrGluLysMetCysAlaGlyIleProGluGluGlyValAspThrCys 381
 DB 1304 GGGAGAGTACCCAG 1363
 QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
 DB 1364 CAGGTGACAGT 1423
 QY 402 ValSerTrpGlyTyrGlyCysGlyIysProSerThrProGlyValIleThrLysValSer 421
 DB 1424 GTTACGTGGGCTTATGCTGT 1483
 QY 422 AlaTyrLeuAsnTrpIleTyrAsnValIleTrpAsnValIleGluLeu 435
 DB 1484 GCTTATCTCACTGAGATCTCAATGTCTGAAAGGCTGAGCTG 1525

RESULT 11
 US-10-956-157-2292
 ; Sequence 2292, Application US/10956157
 ; Publication No. US20050118625A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wyeth
 ; APPLICANT: Mounts, William
 ; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
 ; FILE REFERENCE: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
 ; CURRENT APPLICATION NUMBER: US/10/956,157
 ; NUMBER OF SEQ ID NOS: 319805
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 2292
 ; LENGTH: 2104
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-956-157-2292

Alignment Scores:
 Pred. No.: 1,566-282 Length: 2104
 Score: 2337.00 Matches: 434
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 99.79% Indels: 0
 DB: 9 Gaps: 0

US-10-803-530-2 (1-435) x US-10-956-157-2292 (1-2104)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
 DB 242 GATCCTGACAGTGAATCAACCTCTGAACAGCTCCGATGCAAAACCCCTGGCAACCCCGT 301
 QY 22 IleProMetGluThrPheArgIysValGlyIleProIleIleIleIleAlaLeuLeuSerLeu 41
 DB 302 ATCCCAAG 361
 QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
 DB 362 GCGAGTATCATCATGT 421
 QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
 DB 422 TGGGGAGAGCTTCCACTTCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
 QY 82 ProLeuGlyGluAspGluGluHisCysValIysSerPheProGluGlyProAlaValAla 101
 DB 482 CCCTTGGGGAG 541
 QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsnTrp 121
 DB 542 GTCCGCTCTCCAG 601

PRIOR APPLICATION NUMBER: US 60/315,287
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/372,246
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 163
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 88
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-173-999-88

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
Gaps: 0
DB:

US-10-803-530-2 (1-435) x US-10-173-999-88 (1-1314)

Qy 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValysProLeuArgLysProArg 21
Db 10 GATCCTGACAGTGAACCTCTGAAAGCCTCGATGCAAACTCGGCAAACTCGT 69
Qy 22 ILeProMetGlnThrPheArgLysValGlyIleProIleIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCATGAGACCTTCAGAAAGGTGGGAGATCCCATCATCATATGCACTACTGAGCTG 129
Qy 42 AlAserIleIleIleValIleValIleuIleuValIleLeuAspLysTyrPheLeu 61
Db 130 GCGAGTATCATCATTTGTGTGTCTCTCATCAAGTGATTCGATTAATACTACTCTCTC 189
Qy 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuGlyAspGlyGlnLeuAspCys 81
Db 190 TGGCGGCGACCTCTCCACTTCATCCGAGAAAGCAGCTGTGTGACGGAGAGCTGACTGT 249
Qy 82 ProLeuGlyGlnAspGlnGlnHisCysValysSerPheProGlnGlyProAlaValAla 101
Db 250 CCTTGGGGAGAGAGAGAGACAGCTGTGCAAGACCTTCCCGAAAGGCGCTGCACTGCA 309
Qy 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAAAGACCGATCCACTGCGAGGTGCTGAGCTCGGCCACAGGAACTGG 369
Qy 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTGTGCTGTTTGGACACTTCACAGAGCTCTCGTGAGACAGCCGTGTAGGCAATG 429
Qy 142 GlyTyrSerSerLysProThrPheArgAlaValGluIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCCTACACAGCAAAACCACTTTCAGAGCTGTGGAGATTGGCCAGACAGGATCTGAGT 489
Qy 162 ValValGluIleThrGlnAsnSerGlnGlnLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGGAATCAGAAAAAGCCAGAGGCTTCGATGCGGAAGCTCAAGTGGGCGCTGT 549
Qy 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuThrPro 201
Db 550 CTCTAGAGCTCCCTGCTGCTCCCTGCACTGTCTTGCTGTGGGAAGAGCTGAAGACCCC 609
Qy 202 ArgValValGlyGlyGlnGlnAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGGTGGGGAGAGGCGCTCTGTGATTCCTTGCGCTTGCGAGGTCAATCCAG 669
Qy 222 TyrAspLysGlnHisValCysGlyGlySerIleLeuAspProHisTrpValLeuThrAla 241
Db 670 TACGCAAAACAGACAGCTGTGTGAGGAGACATCTGGACCCCACTGGGTCTTCAAGGGA 729
Qy 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpValArgAlaGlySerAsp 261

Db 730 GCCCAGCTCTTCAGAAACATACGATGTCTTCAACTGGAAGGTGGCGGCTGAC 789
Qy 262 LysLeuGlySerPheProSerLeuAlaValAlaIleIleIleIleIleGluPheAsnPro 281
Db 790 AAATGGGACAGCTTCCATCTCCCTGGCTGTGGCAAGATCATCATCATTTGAATCAACCCC 849
Qy 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAAAGACATGACATGCGCTCATGAAGCTGCAAGTCCCATCATCTTCTCA 909
Qy 302 GlyThrValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAAGGCCCATCTGCTGCTCCCTTCTTGATGAGAGCTCATCCAGCCACCCA 969
Qy 322 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 970 CTCTGTGATCATTTGATGTGGGCTTTTACAAAGAGATGAGGAAAGATGTGACATACCTG 1029
Qy 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGCAAGCGTCAAGTCCAGGTGATTCAGACACAGGTGCAATGCAAGATGCTTACAG 1089
Qy 362 GlyGluValThrGlnLysMetMetCysAlaGlyIleProGlnGlyValAspThrCys 381
Db 1090 GGGGAAGTCAACGAAAGATGATGTGTGCAAGCATCCCGAAAGGGGTGTGACACTGC 1149
Qy 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGTGACAGTGGGGGCCCTGATGATCAATCTGACCAAGTGTGTGTGGGCTATC 1209
Qy 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGTGTGTGGGGGCCGAGCAACCCAGAGATATACCAAGGTCTCA 1269
Qy 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 435
Db 1270 GCCTATCTCACTGATCTACAAATGTCTGGAAGGTGAGCTG 1311

RESULT 10
US-09-851-588-7
Sequence 7, Application US/09851588
Patent No. US20020042067A1
GENERAL INFORMATION:
APPLICANT: Mack, David
APPLICANT: Giesh, Kurt C.
APPLICANT: Wilson, Keith E.
TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER, COMPOSITIONS, AND
FILE REFERENCE: A-68829-1/DJB/JCD/AMS
CURRENT APPLICATION NUMBER: US/09/851,588
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/642,252
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: US 09/656,002
PRIOR FILING DATE: 2000-09-06
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 2081
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (215)..(1528)
OTHER INFORMATION:
US-09-851-588-7

Alignment Scores:
Pred. No.: 1,54e-282 Length: 2081
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 979
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-979

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-979 (1-1314)

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgLysProArg 21
Db 10 GATCTCAACATGATCACTCAACCTGTAACAGCTCGATGTCAACCTCGCGCAACCCCGT 69
QY 22 IleProMetGlnThrPheArgLysValGlyIleProIleIleIleLeuLeuSerLeu 41
Db 70 ATCCCATGAGAGCTTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCTTG 129
QY 42 AlAserIleIleIleValIleValIleValIleValIleValIleValIleValIleVal 61
Db 130 GCGAGTATCATCTGTGTGTTGTTCTTCATCAAGTATTCGTGATTAATCACTTCTTC 189
QY 62 CyeGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCys 81
Db 190 TGGGGAGAGGCTCTCCACTTCATCCGAGGAGCACTGTGTGACGAGAGCTGAGCTGT 249
QY 82 ProLeuGlyGlnAspGlnGlnIleCysValIysSerPheProGlnGlyProAlaValAla 101
Db 250 CCTTGTGGGAGAGCAAGAGAGCACTGTGTCAAGAGCTTCCCGAGAGGCTGCAATGGCA 309
QY 102 ValArgLeuSerIysAspArgSerThrIleGlnIleValIleAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAGAGAGCACTTCACATGACAGGCTGAGCTCGGCCACAGGGAACTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgIleMet 141
Db 370 TTTCTGCTGTTTGCAGCAACTTCACAGAGCTTCGCTGAGACAGCCGTGAGGAGATG 429
QY 142 GlyTyrSerSerIysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGAGCAAACTTCATTCAGAGCTGTGAGATTGGCCAGACAGAGATCTGAT 489

QY 162 ValValGlnIleThrGlnAsnSerGlnGlnLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATCAACAGAAACAGCAGAGCTTCGATCGGAATCAAGTGGGCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuIysThrPro 201
Db 550 CTCTCAGGCTCCCTGTGTTCCCTCGCACTGTCTTGGCTGTGGAGAGCTTGAAGACCC 609
QY 202 ArgValAlaGlyGlyGlnGlnAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGTGGGAGAGAGGCTTGTGATTTCTTGACTTGGCAGGTGAGATCCAG 669
QY 222 TyrAspLysGlnHisValCysGlyIleSerIleLeuAspProHisTrpValIleThrAla 241
Db 670 TACGACAAACAGACAGCTGTGTGAGGAGCACTCGAACCCCACTGGGCTTCACGGCA 729
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAAGAAACATACCATGATGTTCACACTGGAAGTGGGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaIleIleIleIleIleGlnPheAsnPro 281
Db 790 AAACCTGGGAGCTTCCATCCCTGCTGTGGCAAGATCATCATTTGAATTCAAACCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGAACATGACATGATGCTCATGAAAGCTGCACTGCTCACTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnGlnLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAAGCCCATGTCTGTCTCTTGTATGAGAGCTCACTCCAGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTrpGlyPheThrIleGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 970 CTCTGATCATTTGATGAGGCTTTTACAGAGAGATGAGGAGAGATGTCTGACATACTG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGACAGGCTGATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1089
QY 362 GlyGlnValIleThrGlnLysMetMetCysAlaGlyIleProGlnGlyGlyValAspThrCys 381
Db 1090 GGGAGAGTCAACGAGAGATGATGTGACAGCATCCCGAGAGGCTGTGACACTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
Db 1150 CAGGCTGACATGTGTGGGCTTGTATGTATCAATCTGACCAAGTGTGTGGGATC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGCTCGGGGGGCGCGAGCAACCCAGAGATTAACCAAGAGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 435
Db 1270 GCTTATCTCAACTGATCTCAATGTCTGGAAGCTGAGCTG 1311

RESULT 9
US-10-173-999-88
Sequence 88, Application US/10173999
Publication No. US20040005563A1
GENERAL INFORMATION:
APPLICANT: Mack, David H.
APPLICANT: Gish, Kurt C.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
FILE REFERENCE: 018501-002420US
CURRENT APPLICATION NUMBER: US/10/173,999
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: US 60/299,234
PRIOR FILING DATE: 2001-06-18

```

CURRENT APPLICATION NUMBER: US/10/295,027
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2002-11-13
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,354
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,353
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 830
LENGTH: 1314
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-830

```

Alignment Scores:	
Pred. No.:	7.58e-263
Score:	2337.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	99.79%
DB:	6
Length:	1
Matches:	4
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-10-803-530-2 (1-435) x US-10-295-027-830 (1-1314)

QY	2	ASPPROASPSSPASPGLNPROLEUANSERLEUASPVALIYSPROLEUAQGLYSPROARG	21
Db	10	GATCCCTGACAGGATCACTCTGAAACAGCTTCGATGTCAAAACCCCTGCGAATCCCGGT	69
QY	22	ILEPROMETGLUTHIPHEARGLYSVALGILYILEPROILELLEILEALEUENSERLEU	41
Db	70	ATCCCATGAGACCCCTTCAGAAAGGTGGGGATCCCATCATCATADGACCATCTGAGGCTTG	129
QY	42	ALASERILELELEVALVALLEULEULEYSVALILEUENAPLYSTYTRYRPHLEU	61
Db	130	GGAATATCATCATCTGTGGTTGGTCCCTCATCAAGAGTAATTCGTGATTAATCTCATCTTC	189
QY	62	CYSGIYGLNPROLEUHSIPHEILEPROARGLYSGLILEUCYASAPGILYULAUASP	81
Db	190	TCCGGGACACCTCTCCACTTCATCCGAGAGACAGCTGTGTACGAGAGAGCTGGACTGT	248
QY	82	PROLEUGILYULASPGLUNHISCYVALYSSERPHBROGLUGIYPROALAVAJALA	101
Db	250	CCCTTGGGGAGGACGAGGAGCACTGTGTCAAGACTTCCCGAAGGGCTCGCAGTGGCA	309
QY	102	VALARGLEUSERLYSAPARGSERTRHLEUGLINVALLEUASPERALATHIRGIYASERTTP	121
Db	310	GTCCGCTCTCCAGGACCGATCACTGACAGTGCTGGACTCGGCGACAGGGAACTGG	365
QY	122	PHASERIALCYSPHEASPSNPHERTRGUALALEUJAGLUTHRALACYARGGLINMET	141
Db	370	TTCTGTGCGCTGTTTGCACAACTTCACAGAGCTCTGCTGAGACAGCTGTGAGGCAGATG	422
QY	142	GLIYTRYSERSESLYSPROTHRPHARGALAVAGLULLEGLYPROAPGILNAPLEUASP	161
Db	430	GGCTACAGACGAAACCCACTTTCAGAGCTGTGGAGTTGGCCCAACAGAGATCTGAT	489

QY	162	ValValGluIleThrGlnAsnSerGlnGluLeuAqMveLrGAsnSerSerGlyProCys	181
Db	490	GTTGTGAAATCACAGAAAACAGCCAGAGCTTCCGATGCGGAATCAATGAGCCCTGT	549
QY	182	LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuLysThrPro	201
Db	550	CTCTCAGGCTCCCTGCTCTCCCTGCACTGTCTTGCTGTGGGAAAGAGCTGAAGACCCCC	609
QY	202	ArgValValGlyGlyGlnGluAlaSerValAspSerTrrProTrrGlnValSerIleGln	221
Db	610	CGTGTGTGGGTGGGAGAGAGCCCTCTGTGGATTTCTTGGCCCTTGGCAGAGGTCAACAC	669
QY	222	TyrAspLysGlnHisValCysGlyLysSerIleLeuAspProHisTrrValLeuThrAla	241
Db	670	TACGCAAAACAGACAGTCTGTGGAGGAGACATCTTGACCCCACTGGAGTCTTCACGCCA	729
QY	242	AlaHisCysPheArgLysHisTrrAspValPheAsnTrrLysValArgAlaGlySerAsp	261
Db	730	GCCCACTGCTTCAGGAAACATACGATGTTCCACTGGAAGGTGGGCGAGGCTCAGAC	789
QY	262	LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleIleGluPheAsnPro	281
Db	790	AAACTGGGACAGTCTCCCATCCCTGGCTGTGGCCCAAGATCATCATCAATTAATCAACCCC	849
QY	282	MetTrrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer	301
Db	850	ATGTACCCCAAGAACATGACATGCCCTCCATGAAGATGACAGTCCCACTCACTTCTCA	909
QY	302	GlyThrValArgProIleCysLeuProPhePheAspGlnGluLeuThrProAlaThrPro	321
Db	910	GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGAGCTCACTCAGCCACCCCA	969
QY	322	LeuTrrIleIleIleGlyTrrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu	341
Db	970	CTCTGATCATTTGGATGGGGCTTTTACGAAGCAAGAAATGAGGGAAAGATGTCTGACATAC	1020
QY	342	LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTrrGln	361
Db	1030	CTGCAGGCGCTCAGTCCAGGTCAATTGACAGCACCGGTGCMAATGCAAGACATGGGTACAG	1080
QY	362	GlyGluValAlaThrGlnLysMetMetCysAlaGlyIleProGlnGlyGlyValAspThrCys	381
Db	1090	GGGGAAGTCAACGAGAAATGATGTGTGCAGGCAATCCCGAAGGGGTGTGACACTCTGC	1140
QY	382	GlnGlyAspSerSerGlyGlyProLeuMetTrrGlnSerAspGlnTrrHisValValGlyIle	401
Db	1150	CAGGATGACAGTGGTGGGCCCTCGATGATCAATCATGACAGTGGCATGTGGTGGCATC	1200
QY	402	ValSerTrrGlyTrrGlyCysGlyGlyProSerThrProGlyValTrrLysValSer	421
Db	1210	GTTACTGTGGGCTATGTGGCTGCGGGGGGCCGAGACCCCGAGGTATACACCAAGTCTCA	1260
QY	422	AlaTrrLeuAsnTrrIleTrrAsnValTrrLysAlaGluLeu	435
Db	1270	GCCTATCTCACTGATTCATCAATGTCTGAAGGCTGAGCTG	1311
RESULT 8			
US-10-295-027-979			
/ Sequence 979, Application US/10295027			
/ Publication No. US20030232350A1			
/ GENERAL INFORMATION:			
/ APPLICANT: Afar, Daniel			
/ APPLICANT: Aziz, Natasha			
/ APPLICANT: Ginsberg, Wendy M.			
/ APPLICANT: Gish, Kurt C.			
/ APPLICANT: Glyme, Richard			
/ APPLICANT: Hevesi, Peter A.			
/ APPLICANT: Mack, David H.			
/ APPLICANT: Murray, Richard			
/ APPLICANT: Watson, Susan R.			
/ APPLICANT: Eos Biotechnology, Inc.			
/ TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and			
/ TITLE OF INVENTION: Methods of Screening for Modulators of Cancer			

```

; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO: 790
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-790
```

```

Alignment Scores:
Pred. No.: 7,58e-283
Score: 2337.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 99.79%
Length: 1314
Matches: 434
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0
US-10-803-530-2 (1-435) x US-10-295-027-790 (1-1314)
```

```

QY 2 AspProAspSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgLysProArg 21
Db 10 GATCTCTACAGATGATCAACTCTGAGACAGCTCGATGTCAACCCCTGGCAACCCCT 69
QY 22 IleProMetGlnThrPheArgLysValGlyIleProIleIleIleLeuLeuSerLeu 41
Db 70 ATCCCATGAGAGCTTTCAGAAAGGTGGGATCCCATCATCATCATGACCTAGAGCCG 129
QY 42 AlAserIleIleIleValIleValIleValIleValIleValIleValIleValIleVal 61
Db 130 GCGAGTATCATCTGTGTGCTGCTCTCATCAAGGTGATTTCTGGATTAATTAATTAATTA 189
QY 62 CysGlyGlnProLeuHisPheIleProArgLysGlnLeuCysAspGlyGlnLeuAspCys 81
Db 190 TGGGGGAGAGCTCTCCACTTCATCCCGAGAGAGAGCTGTGTGACGAGAGCTGAGCTGT 249
QY 82 ProLeuGlyGlnAspGlnGlnHisCysValIysSerPheProGlnGlyProAlaValAla 101
Db 250 CCTTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 102 ValArgLeuSerLysAspArgSerThrLeuGlnValIleLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
QY 142 GlyTyrSerSerLysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp 161
Db 430 GGCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
QY 162 ValValGlnIleThrGlnAsnSerGlnIleuArgMetArgAsnSerSerGlyProCys 181
```

```

Db 490 GTTGTGAATTCACAGAAACAGCCAGAGCTTCGATCGGAGACTCAAGTGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyLysSerLeuThrPro 201
Db 550 CTCTCAGGCTCCTGTGTCTCCCTGACACTGTCTTGTCTGTGGAGAGAGCTGAGAGAGAG 609
QY 202 ArgValAlaGlyGlyGlnGlnIleAspSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
QY 222 TyrAspLysGlnHisValAlaCysGlyGlySerIleLeuAspProHisTrpValIleuThrAla 241
Db 670 TACGACAAACAGACGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
QY 242 AlaHisCysPheArgLysHisThrAspValPheAsnTrpLysValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAAGAAACATACCATGTGTTCATCTGAAAGTGTGGGAGAGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaLysIleIleIleGlnPheAsnPro 281
Db 790 AAAGTGGGAGCTTCCCATCTCGCTGTGGCAAGATCATCATCATTTGATTCACACCC 849
QY 282 MetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 309
QY 302 GlyThrValArgProIleCysLeuProPhePheAspGlnIleuThrProAlaThrPro 321
Db 910 GGCACAGTCAGAGCCCATGTGTGCCCTTCTTGTATGAGAGCTCATCTCCAGCCCA 369
QY 332 LeuTrpIleIleGlyTrpGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIleLeu 341
Db 970 CTCTGATCATTTGATGTGGGCTTTTACAGACAGATGAGAGAGAGATGTGACATAC 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGACAGGCTGATAGTCCAGGTCATTTGACAGACACAGGAGCAATCAGACGATGGTAC 1089
QY 362 GlyGlnValThrGlnLysMetCysAlaGlyIleProGlnGlyGlyValAspThrCys 381
Db 1090 GGGAGATGACCCGAGAGAGATGTGTGACAGCATCCCGAAGGGGTGTGGACACTGC 1149
QY 382 GlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTrpHisValAlaGlyIle 401
Db 1150 CAGGTATACAGTGTGTGGCCCTGATGTACATTCGACACAGTGCATGTGGTGGC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyGlyProSerThrProGlyValTyrThrLysValSer 421
Db 1210 GTTAGCTGGGGCTATGTGCTGCGGGGGCCGAGAGACCCCGAGAGTATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGlnLeu 435
Db 1270 GCCTATCTCAATCGATCTACAAATGTCTGMAAGGCTGAGCTG 1311

RESULT 7
US-10-295-027-830
; Sequence 830, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natsaba
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
```

```
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 778
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-778

Alignment Scores:
Pred. No.: 7,586-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-778 (1-1314)
QY 2 AspProApeSerAspGlnProLeuAsnSerLeuAspValIysProLeuArgIysProArg 21
DB 10 GATCTGTGACGTGATCAACCTCTGAAACAGCTTCGATGCAAAACCTTCGGCAAAACCTCGT 69
QY 22 IleProMetGluThrPheArgIysValGlyIleProIleIleIleIleValLeuLeuSerLeu 41
DB 70 ATCCCAATGAGAACCTTCGAAAGGTGGGAGATCCCATCATCATATGACACTAGTGAGCTG 129
QY 42 AlaSerIleIleIleValIleValIleValIleValIleValIleValIleValIleValIle 61
DB 130 GCGAGTATCATCATTTGGTGTCTCTCATCAAGGTGATTCGATAAATACTACTTCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysValAspGlyGlnLeuAspCys 81
DB 190 TCGGGGAGCCCTCTCCACTTCATCCCGAGGAAGAGCTGTGACGAGAGCTGAGCTGT 249
QY 82 ProLeuGlyGlnAspGlnGlnIleCysValIlysSerPheProGlnGlyProAlaValAla 101
DB 250 CCTTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValIleAspSerAlaThrGlyAsnTrp 121
DB 310 GTCCGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGlnThrAlaCysArgGlnMet 141
DB 370 TTCTGTGCTGTTTGCACAACTTCAAGAACTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAG 429
QY 142 GlyTyrSerSerIysProThrPheArgAlaValGlnIleGlyProAspGlnAspLeuAsp 161
DB 430 GGCTACAGAGCAAAACCTTTCAGAGCTGTGGAGATGGCCGAGAGAGAGAGAGAGAGAGAGAG 489
QY 162 ValValGlnIleThrGlnAsnSerGlnIleLeuArgMetArgAsnSerSerGlyProCys 181
```

```
DB 490 GTTGTGAATTCACAGAAAACAGCAGAGCTTCGATGGGAACTCAAGTGGCCCTGT 549
QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
DB 550 CTCACAGCTCCCTGGGTCTCTCCCTGCACTGTCTTCCCTGTGGAGAGAGAGAGAGAGAGAG 609
QY 202 ArgValIleGlyIleGlnIleAlaSerValAspSerTrpProTyrGlnValSerIleGln 221
DB 610 CGTGTGGGTGGGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValLeuThrAla 241
DB 670 TACACAAACAGCAGCTGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
QY 242 AlaHisCysPheArgIysHisIleThrAspValPheAsnTrpIysValAspAlaGlySerAsp 261
DB 730 GCCACGCTTCAGAGAAACATACCGAGTGTTCACATGAGAGAGAGAGAGAGAGAGAGAGAG 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleIleIleGlnPheAsnPro 281
DB 790 AAACCTGGGAGCTTCCTCATCTCTGTGTGGCCAGATCATCATCATTTGAATTCACACCC 849
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
DB 850 ATGTACCCCAAAAGACATGACATGCGCCTCATGAAGCTGAGAGTCCACATCATTTCTCA 909
QY 302 GlyThrValArgProIleCysLeuProPheAspGlnIleLeuThrProAlaThrPro 321
DB 910 GGCACAGTCAAGGCCCATCTGTCTCTCTTTATGAGAGAGTCACTCCAGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTyrGlyPheThrIleGlnAsnGlyIysMetSerAspIleLeu 341
DB 970 CTGTGATCATTTGATGGGCTTTTACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
DB 1030 CTGCAAGCGCTGATGCTCAGATTCATGACAGCAGCAGCTGCAATGACAGAGAGCTGACAG 1089
QY 362 GlyIleValThrGlnIysMetMetCysAlaGlyIleProGlnIleGlyIleAspThrCys 381
DB 1090 GGGAGAGTCAACGAGAGAGATGATGTGACAGAGATCCCGAGAGGGGGTGGACACCTGC 1149
QY 382 GlnGlyAspSerGlyIleProLeuMetTyrGlnSerAspGlnTrpHisValIleGlyIle 401
DB 1150 CAGGATGACAGTGTGGGCCCCCTGATGACCAATCTGACCGTGGAGATGTGTGGGATC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIleProSerThrProGlnIleValIleValSer 421
DB 1210 GTTAGCTGGGGCTATGGGTGGGGGCCCGAGCAGCCCGAGAGATATACCAAGAGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpIysAlaGlnLeu 435
DB 1270 GCCTATCTCAACTGAGATTCATATGTCTGAAAGGCTGAGCTG 1311

RESULT 6
US-10-295-027-790
; Sequence 790. Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295, 027
```



```

; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-133

Alignment Scores:
Pred. No.: 7,58e-283 Length: 1314
Score: 2337.00 Matches: 434
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.79% Indels: 0
DB: 6 Gaps: 0

US-10-803-530-2 (1-435) x US-10-295-027-133 (1-1314)
QY 2 AspProAspSerAspGlnProLeuAenSerLeuAspValIysProLeuArgIysProArg 21
Db 10 GATCCTGACGATGATCACTCTGACAGCCTGATGTCACAAACCCCTGGCAAAACCCCTGT 69
QY 22 IleProMetGluThrPheArgIysValIleProIleIleIleAlaLeuLeuSerLeu 41
Db 70 ATCCCATGAGAGCCTTCAGAAAGTGGGAGTCCCATCATCATATACATACATGAGCTG 129
QY 42 AlaSerIleIleIleIleValIleValIleValIleValIleValIleValIleValIle 61
Db 130 GCGAGTATCATCATTTGTGTGTGTCTCATCATCAAGTATTCGTGATTAATTAATTCTCTC 189
QY 62 CysGlyGlnProLeuHisPheIleProArgIysGlnLeuCysAspGlyGluLeuAspCys 81
Db 190 TGGGGGCAAGCTCTTCCACTTTCATCCGAGAGAGCAGCTGTGTGACGAGAGCTGACTGT 249
QY 82 ProLeuGlyGluAspGluGlnIleHisCysValIysSerPheProGluGlyProAlaValAla 101
Db 250 CCTTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
QY 102 ValArgLeuSerIysAspArgSerThrLeuGlnValLeuAspSerAlaThrGlyAsnTrp 121
Db 310 GTCCGCTCTCCAAAGGACGATCCACATGACAGGTGCTGACTCGGCCACAGGAGAACTGG 369
QY 122 PheSerAlaCysPheAspAsnPheThrGlnAlaLeuAlaGluThrAlaCysArgGlnMet 141
Db 370 TTCTCTGCTGTTTGCACAACTTCACAGAACTCTCGCGAACAACAGCTGTATGAGCAGATG 429
QY 142 GlyTyrSerSerIysProThrPheArgAlaValIleGlyProAspGluAspLeuAsp 161
Db 430 GGCCTACAGAGCAAAACCCACTTTCAGAGCTGTGGAGATTTGGCCCAACAGAGATCTGAT 489
QY 162 ValValGluIleThrGluAsnSerGlnGluLeuArgMetArgAsnSerSerGlyProCys 181
Db 490 GTTGTGAATCACAAGAAACAGACGAGAGCTTCCCATGCGGAATCAAGTGGGCGCTGT 549
```

```

QY 182 LeuSerGlySerLeuValSerLeuHisCysLeuAlaCysGlyIysSerLeuIysThrPro 201
Db 550 CTCTGAGGCTCCCTGTCTCCCTGACAGTGTCTGTGCTGGGAGAGAGCTGAAAGCCCC 609
QY 202 ArgValValGlyGlyGluGlnAlaSerValAspSerTrpProTrpGlnValSerIleGln 221
Db 610 CGTGTGTGGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
QY 222 TyrAspIysGlnHisValCysGlyIysSerIleLeuAspProHisTrpValIleThrAla 241
Db 670 TACGACAAACAGACAGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 729
QY 242 AlaHisCysPheArgIysHisThrAspValPheAsnTrpValArgAlaGlySerAsp 261
Db 730 GCCCACTGCTTCAGAAACATACCATATGTTCAACTGAAAGTGTGGGAGGCTCAGAC 789
QY 262 LysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleIleGlnPheAsnPro 281
Db 790 AAACCTGGGAGCTTCCCATCCCTGGCTGTGGCAAGATCATCATTTGAATTCAACCC 849
QY 282 MetTyrProIysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPheSer 301
Db 850 ATGTACCCCAAGACAAAGACATGATGCTCCATGAAAGCTGCACTTCCACTCATCTTCA 909
QY 302 GlyThrValArgProIleCysLeuProPheAspGlnGluLeuThrProAlaThrPro 321
Db 910 GGCACAGTCAGGCCCATCTGTCTGCCCTTCTTGTATGAGAGAGTCACTCCAGCCACCCCA 969
QY 322 LeuTrpIleIleGlyTrpGlyPheThrIleGlnAsnGlyIysLysMetSerAspIleLeu 341
Db 970 CTCTGATCATTTGATGTGGGCTTTTACAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1029
QY 342 LeuGlnAlaSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyrGln 361
Db 1030 CTGACAGGCTGATGATCCAGATGATGATGACAGACAGACAGGTCGATGACAGATCCATCAG 1089
QY 362 GlyIleValIleThrGlyLysMetMetCysAlaGlyIleProGluGlyIysValAspThrCys 381
Db 1090 GGGGAAGTCACCGAAGAGATATGTGTGACAGGATCCCGAAGGGGTGTGGACACTGC 1149
QY 382 GlnGlyAspSerGlyIysProLeuMetTyrGlnSerAspGlnTrpHisValValGlyIle 401
Db 1150 CAGGTGACAGTGTGTGGCTTGTATGATACCAATGTGACAGAGTGTGTGGGAGATC 1209
QY 402 ValSerTrpGlyTyrGlyCysGlyIysProSerThrProGlyValIleThrLysValSer 421
Db 1210 GTTAGCTGGGCTATGCTGCTGGGGGCCAGAGCAACCCAGAGATATACCAAGGTCTCA 1269
QY 422 AlaTyrLeuAsnTrpIleTyrAsnValTrpLysAlaGluLeu 435
Db 1270 GCTATCTCAACTGATCTCAATATGTGTGAAGGCTGAGCTG 1311

RESULT 5
US-10-295-027-778
; Sequence 778, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezl, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Bos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of diagnosis of cancer. Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295, 027
; CURRENT FILING DATE: 2002-11-13
```


CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/11/045,577
FILING DATE: 27-Jan-2005
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/659,151
FILING DATE: 11-Sep-2000
APPLICATION NUMBER: 09/008,271
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2038 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNOT13
CLONE: 137018
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-11-045-577-18

Alignment Scores:
Pred. No.: 1,11e-282 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 10 Gaps: 0

US-10-803-530-2 (1-435) x US-11-045-577-18 (1-2038)

QY 1 MetAASPProASPSeSerPGLInPLeuLnsSerLeuAAspValLySPProLeuAArgLySPPro 20
DB 200 ATGATCTCTGACAGTCAACCTCTGAAACAGCTCGATGTCAACCCCTGGGCAAAACC 259
QY 21 ArgLLePProMetGLInThPheAArgLySValLyLLePProLLeLLeLLeLLeLLeLLeLLe 40
DB 260 CGTATCCCATCGAAGACCTTCGAAAGGTGGGATCCCATCATCATCATCATCATCATCATCAT 319
QY 41 LeuAASerLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeL 60
DB 320 CTGGGAGATCATCATGT 379
QY 61 LeuCyGLyGLInPLeuLnsSerPGLInPLeuLnsSerPGLInPLeuLnsSerPGLInPLeuLns 80
DB 380 CTCTGGGGGAGACCTCTCCATCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 439
QY 81 CysPLeuGLyGLInPLeuLnsSerPGLInPLeuLnsSerPGLInPLeuLnsSerPGLInPLeuLns 100
DB 440 TGTCTCTGGGGGAG 499
QY 101 ALaValAArgLeuSerLySPProThPheAArgLySValLyLLePProLLeLLeLLeLLeLLeLLe 120
DB 500 GCAGTCCGCTCTCCAG 559
QY 121 TrpPheSerAALaCysPheAAspAnPheThArgLLeLLeLLeLLeLLeLLeLLeLLeLLeLLe 140
DB 560 TGTCTCTCTGCTGTCTGCAACAATCTCAAGAAAGCTCTGCTGAGAACAGCTGTAGGAGAG 619
QY 141 MetGLYTYrSerSerLySPProThPheAArgLySValLyLLePProLLeLLeLLeLLeLLeLLe 160
DB 620 ATGGCTACAG 679
QY 161 AspValAArgLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLe 180
DB 680 GATGTGTGTGAATCAAG 739

QY 181 CysLeuSerGLySerLeuValSerLeuLnsCysLeuAALaCysGLyLySPSerLeuLnsThr 200
DB 740 TGTCTCTAGAGCTCTCTGTCT 799
QY 201 ProArgValAArgLyGLyGLInGLInAASerValAAspSerTPProLLeLLeLLeLLeLLeLLe 220
DB 800 CCCCGT 859
QY 221 GLInTyASPLeuSGLInHISValCysGLYGLYSerLLeuAAspProHISTPValLeuThr 240
DB 860 CAGTACACAAACAG 919
QY 241 ALaAAsCysPheAArgLySHISThrASPValPheAAsnTPLySValAArgLLeLLeLLeLLeLLe 260
DB 920 GCAGCCCACTGCTCTCAAGAAACATACGATGTGTCAACTGAGAGAGAGAGAGAGAGAGAGAG 979
QY 261 AspLyLeuGLYSerPheProSerLeuAALaValAArgLLeLLeLLeLLeLLeLLeLLeLLeLLe 280
DB 980 GACAAACTGGGAGAGCTTCCCATCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1039
QY 281 ProMetTyTPProLySPAspAnPheLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLe 300
DB 1040 CCCATGTACCCCAAG 1099
QY 301 SerGLYThrValAArgProLLeCysLeuProPhePheAAspGLInLeuThrProAALaThr 320
DB 1100 TCAGGACAGTCAAG 1159
QY 321 ProLeuTPLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLe 340
DB 1160 CCACTCTGATCATTTGATGTGGGCTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1219
QY 341 LeuLeuGLInAAsPValGInValLLeAAsPSerThArgCysAAsnAALaAspAALaTyR 360
DB 1220 CTGCTGAGGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1279
QY 361 GLInGLyGLInValThrgLInLysMetCysAALaLLePProGLInGLYGLYValAAspThr 380
DB 1280 CAGGGAGAGTCAACGAG 1339
QY 381 CysGLInGLYAsPSerGLYGLYProLeuMetTyRGLInSerAAspGLInTPHISValAArgLy 400
DB 1340 TCCAGGAGTGCAGT 1399
QY 401 LLeValSerTPGLYTYrGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY 420
DB 1400 ATCGTGTAGCTGGGCTATAGT 1459
QY 421 SerAALaTyLeuAAsnTPLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLeLLe 435
DB 1460 TCAGCTTATCTCAACTGGATCTCAATGTCTGAGAGAGCTGAGCTG 1504

RESULT 4
US-10-295-027-133
Sequence 133, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Heyez, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733

ATTORNEY/AGENT INFORMATION:
NAME: Mohan-Peterson, Sheela
REGISTRATION NUMBER: 41,201
REFERENCE/DOCKET NUMBER: PF-0458 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2038 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLANOT13
CLONE: 137018
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-968-415-18

Alignment Scores:
Pred. No.: 1,11e-282 Length: 2038
Score: 2338.00 Matches: 434
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.77% Mismatches: 0
Query Match: 99.83% Indels: 0
Gaps: 0
DB: 3

US-10-803-530-2 (1-435) x US-09-968-415-18 (1-2038)

Qy 1 MetAspProAspSerAspGlnProLeuAsnSerLeuAspValIleuAspProLeuArglyLeuThr 20
Db 200 ATGATCTCTGACAGATCACTCTGAAACAGCTCGATGTAACCCCTGGGCAACCC 259
Qy 21 ArgIlePrometGluThrPheArglyValGlyIleProIleIleIleAlaLeuLeuSer 40
Db 260 CGTATCCCAAGAGACCTTCAGAAAGGTGGGATCCCATCATCATGACACTAGC 319
Qy 41 LeuAlaSerIleIleIleValIleValIleuIleuValIleLeuAspLysTyrTrpPhe 60
Db 320 CTGGGAGATCATCATGTGTGTCTCTCATCAAGTGTGATTTGATTAATCACTATTC 379
Qy 61 LeuCyseGlyGlnProLeuHisPheIleProArglyGlnLeuCyseAspGlyGluLeuAsp 80
Db 380 CTCTGGGGGAGCCCTCTCCACTTCATCCCGAAGAGCAGCTGTGACGAGAGCTGAC 439
Qy 81 CysProLeuGlyGluAspGluGluHisCysValIleYserSerPheProGlyProAlaVal 100
Db 440 TGTCCCTTGGGGAG 499
Qy 101 AlaValArgLeuSerLeuAspArgSerThrLeuGlnValIleuAspSerAlaThrGlyAsn 120
Db 500 GCAAGTCCCTCTCCAAAG 559
Qy 121 TrpPheSerAlaCysePheAspAsnSerThrGluAlaLeuAlaGluThrAlaCysArgGln 140
Db 560 TGTCTCTGCTGCTGTTTGCACAACTTCACAGAAAGCTCTGCGTAGAGAGAGAGAGAG 619
Qy 141 MetGlyTyrSerSerLeuProThrPheArgAlaValIleGluIleGlyProAspGlnAspLeu 160
Db 620 ATGGCTCTACAGATG 679
Qy 161 AspValIleGluIleThrGluAsnSerGlnIleuAspMetArgAsnSerSerGlyPro 180
Db 680 GATGTTGTTGAATATACAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 739
Qy 181 CysLeuSerGlySerLeuValSerLeuHisCysLeuAlaCyseGlyLysSerLeuThr 200
Db 740 TGTCTCTACAGCTCTCTGTCTCTCTGACACTGTCTTGGGAGAGAGAGAGAGAGAGAG 799
Qy 201 ProArgValValGlyGlyGluGluAlaSerValAspSerTyrProTyrGlnValSerIle 220
Db 800 CCCCGTGTGTGGGTGGGATC 859

Qy 221 GlnTyrAspLysGlnHisValCyseGlyGlySerIleLeuAspProHisTyrValIleuThr 240
Db 860 CAGTACAGACAAACAGCAGCTGTGTGAGAGAGAGATCTCGAGAGAGAGAGAGAGAGAGAG 919
Qy 241 AlaAlaHisCysePheArglyHisThrAspValPheAsnTyrPheValAlaGlySer 260
Db 920 GCAAGCCCACTGCTTCAGAGAAACAAACGAGATGTGTTCAACTGAGAGAGAGAGAGAGAG 979
Qy 261 AspLysLeuGlySerPheProSerLeuAlaValAlaValIleIleIleGluPheAsn 280
Db 980 GACAACTGGGAG 1039
Qy 281 ProMetTyrProLysAspAsnAspIleAlaLeuMetLysLeuGlnPheProLeuThrPhe 300
Db 1040 CCATGTATCCCAAGACAAATGATGATGCTCTCTATGAGAGAGAGAGAGAGAGAGAGAGAG 1059
Qy 301 SerGlyThrValArgProIleCyseLeuProPhePheAspGlnIleuThrProAlaThr 320
Db 1100 TCAGGACAGATCAGAGCCATCTGTCTGCTCTTTATGAGAGAGAGAGAGAGAGAGAGAGAG 1159
Qy 321 ProLeuTyrIleIleGlyTyrGlyPheThrLysGlnAsnGlyGlyLysMetSerAspIle 340
Db 1160 CCACTCTGATCATTTGATGAT 1219
Qy 341 LeuLeuGlnIleAspSerValGlnValIleAspSerThrArgCysAsnAlaAspAspAlaTyr 360
Db 1220 CTGCTGAGAGGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1279
Qy 361 GlnGlyGluValThrGluLysMetMetCysAlaGlyIleProGluGlyGlyValAspThr 380
Db 1280 CAGGGAG 1339
Qy 381 CysGlnGlyAspSerGlyGlyProLeuMetTyrGlnSerAspGlnTyrPheIleValGly 400
Db 1340 TGCCAGAGGTGACATGTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1399
Qy 401 IleValSerTyrGlyTyrGlyCyseGlyGlyProSerThrProGlyValTyrThrVal 420
Db 1400 ATCGTTAGCTGGGGCTATGTGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1459
Qy 421 SerAlaTyrLeuAsnTyrIleTyrAsnValTyrPheValGluLeu 435
Db 1460 TCAGCTATCTCAACTGAGATTCATGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1504

RESULT 2
US-10-180-719-18
Sequence 18, Application US/10180719
Publication No. US20030166246A1
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Hillman, Jennifer L.
Yue, Henry
Guegler, Karl J.
Corley, Neil C.
Tang, Tom Y.
Shah, Purvi
TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/180,719
FILING DATE: 25-Jun-2002

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 6, 2005, 21:57:09 ; Search time 903 Seconds
(without alignment)
3983.587 Million cell updates/sec

Title: US-10-803-530-2
Perfect score: 2242
Sequence: 1 MDPDSQPLNSLDVXPLRKP.....VYTKVSAVLMVYVWKAEL 435

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgnt1/USPTO.spool/p/US10803530/runat_05122005_083217_2238/app_query.fasta_1.563
-DB=Published Applications NA Main -QPM=fastcap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -DIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10803530 @CGN 1.1 1549 @runat_05122005_083217_2238
-NCPU=6 -ICPU=3 -NO MAP -LARGECURY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:*

1:	/cgnt2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgnt2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3:	/cgnt2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
4:	/cgnt2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
5:	/cgnt2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6:	/cgnt2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
7:	/cgnt2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
8:	/cgnt2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
9:	/cgnt2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
10:	/cgnt2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2338	99.8	2038	3	US-09-968-415-18 Sequence 18, Appl
2	2338	99.8	2038	6	US-10-180-719-18 Sequence 18, Appl
3	2338	99.8	2038	10	US-11-045-577-18 Sequence 18, Appl
4	2337	99.8	1314	6	US-10-295-027-133 Sequence 133, App
5	2337	99.8	1314	6	US-10-295-027-778 Sequence 778, App
6	2337	99.8	1314	6	US-10-295-027-790 Sequence 790, App
7	2337	99.8	1314	6	US-10-295-027-830 Sequence 830, App
8	2337	99.8	1314	6	US-10-295-027-979 Sequence 979, App

9	2337	99.8	1314	6	US-10-173-999-88 Sequence 88, Appl
10	2337	99.8	2063	3	US-09-851-588-7 Sequence 7, Appl1
11	2337	99.8	2104	9	US-10-956-157-2292 Sequence 2292, Ap
12	2337	99.8	2104	9	US-10-994-117-2 Sequence 2, Appl1
13	2337	99.8	2104	9	US-10-991-287-2 Sequence 2, Appl1
14	2337	99.8	2307	5	US-10-097-340-317 Sequence 317, App
15	2337	99.8	2307	5	US-10-171-311-217 Sequence 217, App
16	2337	99.8	2307	10	US-11-050-926-317 Sequence 317, App
17	2335	99.7	2121	7	US-10-803-530-1 Sequence 1, Appl1
18	2333	99.6	2590	8	US-10-417-375-139 Sequence 139, App
19	2329	99.4	2165	6	US-10-101-510-634 Sequence 634, App
20	2328	99.4	2627	8	US-10-417-375-141 Sequence 141, App
21	2324	99.2	2070	10	US-11-071-974-2 Sequence 2, Appl1
22	2324	99.2	2070	10	US-11-072-918-2 Sequence 2, Appl1
23	2324	99.2	2079	3	US-09-851-588-5 Sequence 5, Appl1
24	2324	99.2	2079	3	US-09-776-191-71 Sequence 71, Appl1
25	2324	99.2	2079	5	US-10-264-820-22 Sequence 22, Appl1
26	2324	99.2	2079	5	US-10-254-289-1 Sequence 1, Appl1
27	2324	99.2	2079	6	US-10-156-214A-38 Sequence 38, Appl1
28	2319	99.0	2137	3	US-09-776-191-3 Sequence 3, Appl1
29	2319	99.0	2137	6	US-10-156-214A-3 Sequence 3, Appl1
30	2297.5	98.1	2063	3	US-09-888-257A-2 Sequence 2, Appl1
31	2297.5	98.1	2063	3	US-09-946-374-214 Sequence 274, App
32	2297.5	98.1	2063	5	US-10-006-867-111 Sequence 111, App
33	2297.5	98.1	2063	5	US-10-052-586-329 Sequence 329, App
34	2297.5	98.1	2063	5	US-10-063-547-111 Sequence 111, App
35	2297.5	98.1	2063	5	US-10-063-551-111 Sequence 111, App
36	2297.5	98.1	2063	5	US-10-174-581-329 Sequence 329, App
37	2297.5	98.1	2063	5	US-10-176-749-329 Sequence 329, App
38	2297.5	98.1	2063	5	US-10-176-758-329 Sequence 329, App
39	2297.5	98.1	2063	5	US-10-175-737-329 Sequence 329, App
40	2297.5	98.1	2063	5	US-10-063-616-111 Sequence 111, App
41	2297.5	98.1	2063	5	US-10-174-581-329 Sequence 111, App
42	2297.5	98.1	2063	5	US-10-176-483-329 Sequence 329, App
43	2297.5	98.1	2063	5	US-10-176-749-329 Sequence 329, App
44	2297.5	98.1	2063	5	US-10-176-914-329 Sequence 329, App
45	2297.5	98.1	2063	5	US-10-176-915-329 Sequence 329, App
					US-10-063-569-111 Sequence 111, App

ALIGNMENTS

RESULT 1
US-09-968-415-18
; Sequence 18, Application US/09968415
; Publication No. US20020086334A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Hillman, Jennifer L.
; Yue, Henry
; Guegler, Karl J.
; Corley, Neil C.
; Tang, Tom Y.
; Shah, Purvi
; TITLE OF INVENTION: HUMAN PROTEASE MOLECULES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/968, 415
; FILING DATE: 26-Sep-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/659, 151
; FILING DATE: <Unknown>

THIS PAGE BLANK (USPTO)

```

Db      1183  TTCTGTGCTGGCTTCATGAAAGAGTAGAGATTCACTCAAGGAGHTAGTGGGGAGCC 1242
Qy      389  LeuMetTyrGlnSerAsp---GlnTrpHisValValGlyTleValSerTyrGly 407
Db      1243  CATGTTACTGAAAGTGAAGGACCACTTCTTAAGTGAATATATAGCTGGGGTGAAGAG 1302
Qy      408  CysGlyGlyProSerThrProGlyValTyrThrIysValSerAlaTyrLeuMetTrpIle 427
Db      1303  TGTGCAATGAAGGCAATATGGAATATATACCAAGGTATCCGGTATGTCAACTGATT 1362

```

Search completed: December 7, 2005, 01:18:48
 Job time : 256 secs